CANADIAN JOURNAL OF MATHEMATICS

Journal Canadien de Mathématiques

VOL. XI · NO. 2
1959

Published for
THE CANADIAN MATHEMATICAL CONGRESS
by the
University of Toronto Press

The chief languages of the *Journal* are English and French.

Manuscripts for publication in the *Journal* should be sent to the *Editor-in-Chief*, G. F. D. Duff, University of Toronto. Authors are asked to write with a sense of perspective and as clearly as possible, especially in the introduction. Regarding typographical conventions, attention is drawn to the Author's Manual of which a copy will be furnished on request.

All other correspondence should be addressed to the *Managing Editor*, G. de B. Robinson, University of Toronto.

The *Journal* is published quarterly. Subscriptions should be sent to the *Managing Editor*. The price per volume of four numbers is \$8.00. This is reduced to \$4.00 for individual members of recognized Mathematical Societies.

The Canadian Mathematical Congress gratefully acknowledges the assistance of the following towards the cost of publishing this Journal:

University of Alberta	Assumption University
University of British Columbia	Carleton College
Dalhousie University	Ecole Polytechnique
Université Laval	Loyola College
University of Manitoba	McGill University
McMaster University	Université de Montréal
Queen's University	Royal Military College
St. Mary's University	University of Toronto

National Research Council of Canada
and the
American Mathematical Society

AUTHORIZED AS SECOND CLASS MAIL, POST OFFICE DEPARTMENT, OTTAWA

SOME REMARKS ON PRIME FACTORS OF INTEGERS

P. ERDÖS

1. Let $1 < a_1 < a_2 < \ldots$ be a sequence of integers and let $N(x)$ denote the number of a's not exceeding x. If $N(x)/x$ tends to a limit as x tends to infinity we say that the a's have a density. Often one calls it the asymptotic density to distinguish it from the Schnirelmann or arithmetical density. The statement that almost all integers have a certain property will mean that the integers which do not have this property have density 0. Throughout this paper p, q, r will denote primes.

I conjectured for a long time that, if $\epsilon > 0$ is any given number, then almost all integers n have two divisors d_1 and d_2 satisfying

$$d_1 < d_2 < (1 + \epsilon)\, d_1. \tag{1}$$

I proved (**1**, p. 691) that the integers with two divisors satisfying (1) have a density, but I cannot prove that this density has the value 1. However, analogous questions can be asked about the prime divisors of integers and a more complete result is contained in the following theorem.

THEOREM 1. *Let $\epsilon_p > 0$, $\delta_p = \epsilon_p$ if $\epsilon_p < 1$ and $\delta_p = 1$ if $\epsilon_p > 1$. The divergence of $\sum_p \delta_p/p$ is a necessary and sufficient condition that almost all integers should have two prime factors p and q satisfying*

$$p < q < p^{1+\epsilon_p}. \tag{2}$$

From the prime number theorem we have

$$p_n = (1 + o(1))n \log n;$$

thus $\sum_p \epsilon_p p^{-1}$ will diverge if $\epsilon_p = (\log \log p)^{-1}$, but will converge if $\epsilon_p = (\log \log p)^{-1-c}$, for any $c > 0$.

Further, we shall outline a proof of

THEOREM 2. *The density of integers n which have two prime factors p and q satisfying*

$$p < q < p^{1+c/\log\log n}$$

equals $1 - e^{-c}$.

Let $p_1 < p_2 < \ldots < p_k$ be the distinct prime factors of n. Define the real number η_i by $p_i^{\eta_i} = p_{i+1}$. A famous result of Hardy and Ramanujan (**4**) asserts that $k = (1 + o(1)) \log \log n$ for almost all n. I proved (**2**, p.

Received April 23, 1958.

533, Theorem 9) that, for almost all n, the number of η's not exceeding t $(t > 1)$ is

$$(1 + o(1))\left(1 - \frac{1}{t}\right) \log \log n.$$

Theorem 2 can be stated as follows: the density of integers with

$$\min_{1 \leqslant i \leqslant k} \eta_i < 1 + \frac{c}{\log \log n}$$

is $1 - e^{-c}$. By similar methods, we can prove that the density of integers n satisfying

$$\max_{1 \leqslant i \leqslant k} \eta_i > c \log \log n$$

is $1 - \exp[-1/c]$. Further, we can prove that the divergence of $\sum_p \delta_p/p$ $(\delta_p < 1)$ is the necessary and sufficient condition that almost all integers n should have a prime factor p such that $n \equiv 0 \pmod p$, and $n \not\equiv 0 \pmod q$ for all primes with

$$p \leqslant q < p^{\delta_p^{-1}}.$$

We shall not give the proof of these results, since they are similar to those of Theorems 1 and 2.

2. First, we show that the condition of Theorem 1 is necessary. In fact, we show that if $\sum_p \delta_p/p < \infty$, then the upper density of integers having two prime divisors satisfying (2) is less than one. Since $\sum_p \delta_p p^{-1} < \infty$, it is clear that

$$\sum_{\epsilon_p > 1} p^{-1} < \infty.$$

Denote by $b_1 < b_2 < \ldots$ the integers consisting of the primes p satisfying $\epsilon_p \geqslant 1$ and the integers of the form pq, where $\epsilon_p < 1$ and $p < q < p^{1+\epsilon_p}$. Clearly the integers not divisible by any b have no divisor of the form pq satisfying (2). But $\sum b_i^{-1} < \infty$; thus by a well-known and simple argument (**3**, p. 279) one can show that the density of integers divisible by a b is less than one. We really only proved that if $\sum_p \delta_p/p < 1$ then the upper density of integers having a divisor of the form pq satisfying (2) is less than one. In fact it would be quite easy to show that the density in question exists.

Now we prove the sufficiency of Theorem 1. We first show that it will suffice to prove the following

THEOREM 1′. *Let $\epsilon_p < \frac{1}{4}$, $\epsilon_p \to 0$, $\sum_p \epsilon_p/p = \infty$. Then the density of integers n having two prime divisors p and q satisfying*

$$p < q < p^{1+\epsilon_p}$$

is 1.

To deduce the sufficiency of the condition of Theorem 1 from Theorem 1′ it will suffice to show that if $\sum_p \delta_p/p = \infty$ there always exists an $\epsilon_p' \leqslant \epsilon_p$,

$\epsilon_p' < \frac{1}{4}$, $\sum_p \epsilon_p'/p = \infty$. To show this we observe that if $\sum_p \delta_p/p = \infty$ then either there exists a subsequence p_i with

$$\sum_i \epsilon_{p_i} p_i^{-1} = \infty, \qquad \epsilon_{p_i} < \tfrac{1}{4}$$

and then we put

$$\epsilon'_{p_i} = \epsilon_{p_i}, \qquad 1 \leqslant i < \infty,$$

$\epsilon_p' = 0$ if $p \neq p_i$, or for a certain

$$c > \tfrac{1}{8}, \qquad \sum_{\epsilon_p > c} p^{-1} = \infty.$$

But in this case there clearly exists an $\epsilon_p' < \epsilon_p$ such that

$$\epsilon_p' \to 0, \qquad \epsilon_p' < \tfrac{1}{4}, \qquad \sum \frac{\epsilon_p'}{p} = \infty,$$

which completes our proof.

Now we prove Theorem 1′. Put

$$\sum_{p<x} \frac{1}{p} \sum_{p<q<p^{1+\epsilon_p}} \frac{1}{q} = A(x);$$

then, since $\sum_p \epsilon_p/p = \infty$,

$$A(x) \to \infty \quad \text{as} \quad x \to \infty.$$

We have to show that almost all integers have at least one divisor of the form pq, where $p < q < p^{1+\epsilon_p}$. Instead of this we shall prove the stronger result that if $f(n)$ denotes the number of divisors of n of the above form then, for almost all n,

$$f(n) = (1 + o(1))A(n). \tag{3}$$

Or, because of the slow growth of $A(n)$, we shall in fact prove that

$$f(n) = (1 + o(1))A(x), \tag{4}$$

except for $o(x)$ values of $n < x$. It is easy to see that (3) and (4) are equivalent since

$$A(x) - A(x^{\frac{1}{2}}) = \sum_{x^{\frac{1}{2}}<p<x} \frac{1}{p} \sum_{p<q<p^{1+\epsilon_p}} \frac{1}{q} = \sum_{x^{\frac{1}{2}}<p<x} \frac{\epsilon_p + o(1)}{p} = o(1) \tag{5}$$

by the well-known estimate

$$\sum_{p<x} p^{-1} = \log\log x + c_1 + O\left(\frac{1}{\log x}\right).$$

To prove (4) we shall use Turan's method (**6**, pp. 274–6). We have

$$\sum_{n=1}^{x} (f(n) - A(x))^2 = x(A(x))^2 - 2A(x)\sum_{n=1}^{x} f(n) + \sum_{n=1}^{x} f^2(n). \tag{6}$$

Since

$$f(n) = \sum_{\substack{pq \mid n \\ p<q<p^{1+\epsilon_p}}} 1,$$

we may write

$$\sum_{n=1}^{x} f(n) = \sum_{p<x} \sum_{p<q<p^{1+\epsilon_p}} \left[\frac{x}{pq}\right] = x \sum_{p}{}' \sum_{p<q<p^{1+\epsilon_p}}{}' \frac{1}{pq} + O(x), \tag{7}$$

where the dash indicates that $pq < x$.* Now $\epsilon_p < \frac{1}{4}$ implies that for $p < x^{\frac{1}{3}}$, $pq < x$

$$A(x^{\frac{1}{3}}) < \sum_{p}{}' \sum_{p<q<p^{1+\epsilon_p}}{}' \frac{1}{pq} < A(x).$$

Thus from (5),

$$\sum_{n=1}^{x} f(n) = xA(x) + O(x). \tag{8}$$

Similarly,

$$\sum_{n=1}^{x} f^2(n) = \sum_{p<x} \sum_{p<q<p^{1+\epsilon_p}} \left[\frac{x}{pq}\right] + \sum \sum \left[\frac{x}{\{p_1q_1, p_2q_2\}}\right], \tag{9}$$

where in the second sum

$$p_1 < q_1 < p_1^{1+\epsilon_{p_1}}, \qquad p_2 < q_2 < p_2^{1+\epsilon_{p_2}},$$

$p_1q_1 \neq p_2q_2$, and ($\{p_1q_1, p_2q_2\}$ denotes the least common multiple of p_1q_1 and p_2q_2).

The first sum on the right of (9) is $(1 + o(1))xA(x)$. For the second sum we have

$$\sum \sum \left[\frac{x}{\{p_1q_1, p_2q_2\}}\right] = x \sum{}' \sum{}' \frac{1}{\{p_1q_1, p_2q_2\}} + O(x), \tag{10}$$

where the dash indicates that $p_1q_1 \neq p_2q_2$ and $\{p_1q_1, p_2q_2\} < x$. Clearly, from (5), if $p_1 < x^{\frac{1}{4}}$, $p_2 < x^{\frac{1}{4}}$, $\{p_1q_1, p_2q_2\} < x$

$$\sum{}' \sum{}' \frac{1}{\{p_1q_1, p_2q_2\}} > (A(x^{\frac{1}{4}}))^2 + O(1) = (1 + o(1))(A(x))^2. \tag{11}$$

On the other hand, by a simple argument,

$$\sum{}' \sum{}' \frac{1}{\{p_1q_1, p_2q_2\}} < A^2(x) + 4 \sum{}'' \frac{1}{r_1r_2r_3}, \tag{12}$$

where in $\sum''$

$$r_1 < r_2 < r_1^{1+\epsilon_{r_1}}, \qquad r_3 < \max(r_1^{1+\epsilon_{r_1}}, \qquad r_2^{1+\epsilon_{r_2}}),$$

or $r_3 < r_1^2$, and $r_1 < x$. (12) follows from the fact that $r_1r_2r_3 = \{p_1q_1, p_2q_2\}$ has four solutions. Now

---

*Since $p < q$, the equation $pq = \lambda$ has at most one solution p, q and so there are at most x terms in the double sum. Hence the error in omitting the square brackets is at most x.

$$\sum{}'' \frac{1}{r_1 r_2 r_3} < \sum_{p<x} \frac{1}{p} \sum_{p<q<p^{1+\epsilon_p}} \frac{1}{q} \sum_{p<r<p^2} \frac{1}{r} < c\,A(x),$$

hence

$$\sum{}' \sum{}' \frac{1}{\{p_1 q_1, p_2 q_2\}} = (1 + o(1))A^2(x). \tag{13}$$

Thus, by (9), (10), and (13),

$$\sum_{n=1}^{x} f^2(n) = (1 + o(1))x(A(x))^2. \tag{14}$$

Hence from (6), (8), and (14)

$$\sum_{n=1}^{x} (f(n) - A(x))^2 = o(xA^2(x)),$$

which proves that $f(n) = (1 + o(1))A(x)$, except for $o(x)$ values of $n < x$. Thus Theorem 1 is proved.

3. Now we outline the proof of Theorem 2. Denote by $a_1 < a_2 < \dots < a_k \leq x$, the integers not exceeding x of the form pq, where

$$p < q < p^{1+c/\log\log x}.$$

Clearly the a's depend on x and $a_1 \to \infty$ as x tends to infinity. Denote by $N_c(a_1, \dots, a_k; x)$ the number of integers not exceeding x which are not divisible by any of the a_i's. Further, denote by $M_c(x)$ the number of integers $n \leq x$ which do not have two prime factors p and q satisfying

$$p < q < p^{1+c/\log\log n}.$$

We have to prove that

$$M_c(x) = (1 + o(1))e^{-c}x. \tag{15}$$

Clearly

$$M_c(x) \leq N_c(a_1, a_2, \dots, a_k; x),$$

but because of the slow increase of $\log\log n$ it is easy to see that

$$M_c(x) = N_c(a_1, a_2, \dots, a_k; x) + o(x).$$

Thus to prove Theorem 2 it will suffice to show that

$$N_c(a_1, a_2, \dots, a_k; x) = xe^{-c} + o(x). \tag{16}$$

We obtain by a simple sieve process the well-known formula

$$N_c(a_1, a_2, \dots, a_k; x) = x \sum_{l=0}^{k} (-1)^l \sum{}_l,$$

where

$$\sum{}_0 = 1, \quad \text{and} \quad \sum{}_l = \sum \frac{1}{\{a_{i_1}, \dots, a_{i_l}\}},$$

where $i_1, i_2, \ldots, i_l$ runs through all distinct l-tuples from 1 to k. (The curly bracket in the denominator denotes least common multiple.)

By a well-known combinatorial argument*

$$(17)\qquad x\sum_{l=0}^{2t-1}\left((-1)^l\sum\nolimits_l\right) < N(a_1,\ldots,a_k;x) < x\sum_{l=0}^{2t}\left((-1)^l\sum\nolimits_l\right),$$

for every $t > 0$. We evidently have, by a simple computation (the dashes indicate that $p < q < p^{1+c/\log\log x}$ and $pq < x$)

$$(18)\qquad \sum_{i=1}^{k}\frac{1}{a_i} = \sum{}'\frac{1}{p}\sum{}'\frac{1}{q} = \frac{(1+o(1))c}{\log\log x}\sum_{p<x}\frac{1}{p} + o(1) = c + o(1),$$

by the estimate for $\sum_{p<x} p^{-1}$. Further, for every fixed l (the two dashes indicate that

$$\{a_{i_1}, a_{i_2}, \ldots, a_{i_l}\} < x)$$

$$(19)\qquad \sum\nolimits_l = \sum\left[\frac{x}{\{a_{i_1},\ldots,a_{i_l}\}}\right] = x\sum\nolimits_l''\frac{1}{\{a_{i_1},\ldots,a_{i_l}\}} + o(x)$$
$$= x\sum\nolimits_l'' + o(x),$$

since there are only $o(x)$ l-tuples satisfying

$$\{a_{i_1},\ldots,a_{i_l}\} < x.$$

This last statement follows from the fact that the integers

$$\{a_{i_1},\ldots,a_{i_l}\}$$

have at most $2l$ prime factors and, by a well-known theorem of Landau (**5**, Vol. I, pp. 208–11), the number of integers not exceeding x having $2l$ prime factors equals

$$(1+o(1))\frac{x}{\log x}\frac{(\log\log x)^{2l-1}}{(2l-1)!} = o(x),$$

and finally a simple argument shows that the number of solutions of

$$y = \{a_{i_1},\ldots,a_{i_l}\}$$

is less than a constant depending only on l.

Now we outline the proof of

$$(20)\qquad \sum\nolimits_l'' = \frac{c^l}{l!} + o(1).$$

For $l = 1$, (20) follows from (18). For $l > 1$ we can prove (20) by a simple induction process, similar but a bit more complicated than that used in the estimations in Theorem 1. We do not give the details since they are somewhat cumbersome.

---

*This is one of the basic ideas of Brun's method, see for example, Landau *Zahlentheorie*, Vol. 1, Kap. 2.

From (17) and (20) we have

$$N_c(a_1, \ldots, a_k; x) = x \sum_{l=0}^{\infty} \frac{(-1)^l c^l}{l!} + o(x) = x\, e^{-c} + o(x), \tag{21}$$

which is (16).

References

1. P. Erdös, *On the density of some sequences of integers*, Bull. Amer. Math. Soc., *54* (1948), 691.
2. ———, *Some remarks about additive and multiplicative functions*, Bull. Amer. Math. Soc., *52* (1946), 533, theorem 9.
3. ———, *On the density of the abundant numbers*, J. Lond. Math. Soc., *9* (1934), 279.
4. G. H. Hardy and S. Ramanujan, *The normal number of prime factors of n*, Quart. J. Math., *48* (1917), 76–92.
5. E. Landau, *Verteilung der Primzahlen*,
6. P. Turán, *On a theorem of Hardy and Ramanujan*, J. Lond. Math. Soc., *9* (1934), 274–6.

University of Alberta

ON THE REPRESENTATION OF FUNCTIONS AS FOURIER TRANSFORMS

P. G. ROONEY

If $f \in L_p(-\infty, \infty)$, $1 < p \leqslant 2$, then f has a Fourier-Plancherel transform $F \in L_q(-\infty, \infty)$ where $p^{-1} + q^{-1} = 1$. Also if $|x|^{1-2/q} f(x) \in L_q(-\infty, \infty)$, $q \geqslant 2$, then f has a Fourier-Plancherel transform $F \in L_q(-\infty, \infty)$. These results can be found in (**2**, Theorems 74 and 79). In neither case, however, does the collection of transforms cover L_q, except when $p = q = 2$, and in neither case, with the same exception, has the collection of transforms been characterized.

Further, if $f \in L_p(-\infty, \infty)$, $1 < p \leqslant 2$, then its transform F has the property $|x|^{1-2/p} F(x) \in L_p(-\infty, \infty)$ (see **2**, Theorem 80) but, except when $p = 2$, the collection of transforms does not cover the set of functions with this property, and again, except when $p = 2$, the collection of transforms has not been characterized.

Our object here is to find such characterizations, and this is done for the various cases in Theorems 1, 2, and 3 below. This characterization is given in terms of an operator

$$\mathfrak{F}_{k,t}[F] = \frac{(-ik/t)^{k+1}}{(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{\infty} \frac{1}{(x-ik/t)^{k+1}} F(x)dx, \qquad k = 1, 2, \ldots .$$

It transpires that this operator is an inversion operator for the Fourier transform, and its inversion theory will be the subject of another paper.

THEOREM 1. *A necessary and sufficient condition that a function $F \in L_q(-\infty, \infty)$, $q \geqslant 2$, be the Fourier transform of a function in $L_p(-\infty, \infty)$, with $p^{-1} + q^{-1} = 1$, is that there exist a constant M such that*

$$\int_{-\infty}^{\infty} |\mathfrak{F}_{k,t}[F]|^p dt < M, \qquad k = 1, 2, \ldots .$$

Proof of necessity. Suppose F is the Fourier transform of $f \in L_p(-\infty, \infty)$. Now an easy calculation shows that for $k = 1, 2, \ldots,$

$$\frac{1}{(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{\infty} \frac{1}{(x-ik/t)^{k+1}} e^{-ixy} dx = \begin{cases} -(2\pi)^{\frac{1}{2}}(-i)^{k+1} y^k e^{ky/t}/k!, & y < 0, t > 0, \\ (2\pi)^{\frac{1}{2}}(-i)^{k+1} y^k e^{ky/t}/k!, & y > 0, t < 0, \\ 0 \qquad , & yt > 0. \end{cases}$$

Hence, since for each $t \neq 0$ and each $k = 1, 2, \ldots$, $(x - ik/t)^{-(k+1)} \in L_p(-\infty, \infty)$, we have from (**2**, Theorem 75) that

Received April 23, 1958.

$$\mathfrak{F}_{k,t}[F] = \frac{(-ik/t)^{k+1}}{(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{\infty} \frac{1}{(x-ik/t)^{k+1}} F(x)dx$$

$$= \begin{cases} (k/t)^{k+1}(k!)^{-1} \int_0^\infty e^{-ky/t} y^k f(y)dy, & t > 0 \\ (k/|t|)^{k+1}(k!)^{-1} \int_{-\infty}^0 e^{-ky/t}|y|^k f(y)dy, & t < 0. \end{cases}$$

Thus, using Hölder's inequality, we have for $t > 0$

$$|\mathfrak{F}_{k,t}[F]| < (k/t)^{k+1}(k!)^{-1}\left\{\int_0^\infty e^{-ky/t}y^k|f(y)|^p dy\right\}^{1/p}\left\{\int_0^\infty e^{-ky/t}y^k dy\right\}^{1/q}$$
$$= \left\{(k/t)^{k+1}(k!)^{-1}\int_0^\infty e^{-ky/t}|f(y)|^p dy\right\}^{1/p},$$

and consequently,

$$\int_0^\infty |\mathfrak{F}_{k,t}[F]|^p dt < \frac{k^{k+1}}{k!}\int_0^\infty t^{-(k+1)}\, dt \int_0^\infty e^{-ky/t}\, y^k|f(y)|^p dy$$
$$= \frac{k^{k+1}}{k!}\int_0^\infty y^k|f(y)|^p dy \int_0^\infty t^{-(k+1)}e^{-ky/t}dt = \int_0^\infty |f(y)|^p dy.$$

A similar calculation for $t < 0$ shows that

$$\int_{-\infty}^0 |\mathfrak{F}_{k,t}[F]|^p dt < \int_{-\infty}^0 |f(y)|^p dy,$$

and hence

$$\int_{-\infty}^\infty |\mathfrak{F}_{k,t}[F]|^p dt < \int_{-\infty}^\infty |f(y)|^p dy = M.$$

Proof of sufficiency. For $s > 0$ let

$$g_+(s) = -(2\pi)^{-\frac{1}{2}} i \int_{-\infty}^\infty \frac{1}{x - is} F(x)dx,$$

and

$$g_-(s) = (2\pi)^{-\frac{1}{2}} i \int_{-\infty}^\infty \frac{1}{x + is} F(x)dx,$$

and denote by $L_{k,t}$ the Widder-Post inversion operator for the Laplace transformation; that is

$$L_{k,t}[g] = (-1)^k(k/t)^{k+1} g^{(k)}(k/t)/k!, \qquad k = 1, 2, \ldots.$$

Now if $s \geq \delta > 0$, and $k = 1, 2, \ldots$, then

$$|(x \pm is)^{-(k+1)} F(x)| < (x^2 + \delta^2)^{-(k+1)/2}|F(x)| \in L_1(-\infty, \infty),$$

since from Hölder's inequality

$$\int_{-\infty}^\infty (x^2 + \delta^2)^{-(k+1)/2}|F(x)|dx$$
$$< \left\{\int_{-\infty}^\infty (x^2 + \delta^2)^{-p(k+1)/2}dx\right\}^{1/p} \cdot \left\{\int_{-\infty}^\infty |F(x)|^q dx\right\}^{1/q} < \infty.$$

Hence by (**1**, Corollary 39.2), $g_{\pm}(s)$ has derivatives of all orders in $0 < s < \infty$, and these derivatives can be calculated by differentiating under the integral sign. Thus for $t > 0$,

$$L_{k,t}[g_+] = \frac{(-ik/t)^{k+1}}{(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{\infty} \frac{1}{(x - ik/t)^{k+1}} F(x)dx = \mathfrak{F}_{k,t}[F],$$

and

$$L_{k,t}[g_-] = \frac{(ik/t)^{k+1}}{(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{\infty} \frac{1}{(x + ik/t)^{k+1}} F(x)dx = \mathfrak{F}_{k,-t}[F],$$

so that

$$\int_0^\infty |L_{k,t}[g_+]|^p dt = \int_0^\infty |\mathfrak{F}_{k,t}[F]|^p dt < M, \qquad k = 1, 2, \ldots,$$

and

$$\int_0^\infty |L_{k,t}[g_-]|^p dt = \int_0^\infty |\mathfrak{F}_{k,-t}[F]|^p dt < M, \qquad k = 1, 2, \ldots.$$

Further $g_{\pm}(s) \to 0$ as $s \to \infty$. For from Hölder's inequality we have

$$|g_{\pm}(s)| < (2\pi)^{-\frac{1}{2}} \left\{\int_{-\infty}^{\infty} (x^2 + s^2)^{-p/2} dx\right\}^{1/p} \left\{\int_{-\infty}^{\infty} |F(x)|^q dx\right\}^{1/q} = 0(s^{-1/q}).$$

Hence by (**3**, Chapter 7, Theorem 15a) there are functions f_+ and f_- in $L_p(0, \infty)$ such that

$$g_+(s) = \int_0^\infty e^{-st} f_+(t)dt, \qquad s > 0,$$

and

$$g_-(s) = \int_0^\infty e^{-st} f_-(t)dt, \qquad s > 0.$$

Let

$$f(t) = \begin{cases} f_+(t), & t > 0, \\ f_-(-t), & t < 0. \end{cases}$$

Then clearly $f \in L_p(-\infty, \infty)$ and hence by (**2**, Theorem 74) f has a Fourier transform $F^* \in L_q(-\infty, \infty)$. We now show $F = F^*$ a.e.

Let

$$g_+^*(s) = -(2\pi)^{-\frac{1}{2}}\, i \int_{-\infty}^{\infty} \frac{1}{x - is} F^*(x)dx, \qquad s > 0,$$

and

$$g_-^*(s) = (2\pi)^{-\frac{1}{2}}\, i \int_{-\infty}^{\infty} \frac{1}{x + is} F^*(x)dx, \qquad s > 0.$$

Then since for each $s > 0$, $(x - is)^{-1} \in L_p(-\infty, \infty)$, and

$$(2\pi)^{-\frac{1}{2}}(P) \int_{-\infty}^{\infty} \frac{1}{(x - is)} e^{-ixy} dx = \begin{cases} (2\pi)^{\frac{1}{2}}\, i\, e^{sy}, & y < 0, s > 0, \\ -(2\pi)^{\frac{1}{2}}\, i\, e^{sy}, & y > 0, s < 0, \\ 0, & sy > 0, \end{cases}$$

we have from (2, Theorem 75) for $s > 0$,

$$g_+^*(s) = -(2\pi)^{-\frac{1}{2}}\, i \int_{-\infty}^{\infty} \frac{1}{x - is} F^*(x)dx$$
$$= \int_0^{\infty} e^{-sy} f(y)dy = \int_0^{\infty} e^{-sy} f_+(y) = g_+(s),$$

and

$$g_-^*(s) = (2\pi)^{-\frac{1}{2}}\, i \int_{-\infty}^{\infty} \frac{1}{x + is} F^*(x)dx$$
$$= \int_{-\infty}^{0} e^{sy} f(y)dy = \int_0^{\infty} e^{-sy} f_-(y)dy = g_-(s).$$

Consequently, for $s > 0$

$$\int_{-\infty}^{\infty} \frac{1}{x - is} (F(x) - F^*(x))dx = 0$$

and

$$\int_{-\infty}^{\infty} \frac{1}{x + is} (F(x) - F^*(x))dx = 0.$$

Letting $\phi(x) = F(x) - F^*(x)$, the last two equations yield

$$\int_{-\infty}^{\infty} \frac{1}{x + is} \phi(x)dx = 0, \qquad s \neq 0.$$

Then denoting the even and odd parts of ϕ by ϕ_e and ϕ_0 respectively, we have for $s \neq 0$

$$\int_{-\infty}^{\infty} \frac{1}{x + is} \phi_e(x)dx = -\int_{-\infty}^{\infty} \frac{1}{x + is} \phi_0(x)dx.$$

But the function on the left of this equation is an odd function of s while the function on the right is even. Hence each is zero, so that for $s \neq 0$

$$\int_0^{\infty} \frac{1}{x^2 + s^2} \phi_e(x)dx = -\frac{1}{2is}\int_{-\infty}^{\infty} \frac{1}{x + is} \phi_e(x)dx = 0,$$

and

$$\int_0^{\infty} \frac{x}{x^2 + s^2} \phi_0(x)dx = -\tfrac{1}{2}\int_{-\infty}^{\infty} \frac{1}{x + is} \phi_0(x)dx = 0.$$

Thus for each $s > 0$,

$$\int_0^{\infty} \frac{1}{x + s} x^{-\frac{1}{2}} \phi_e(x^{\frac{1}{2}})dx = 2\int_0^{\infty} \frac{1}{x^2 + s} \phi_e(x)dx = 0,$$

and

$$\int_0^{\infty} \frac{1}{x + s} \phi_0(x^{\frac{1}{2}})dx = 2\int_0^{\infty} \frac{x}{x^2 + s} \phi_0(x)dx = 0,$$

and hence by the uniqueness theorem for the Stieltjes transformation (3, chapter 8, Theorem 5a) ϕ_e and ϕ_0 are zero almost everywhere. Thus ϕ is zero

almost everywhere so that $F = F^*$ almost everywhere, and F has the prescribed representation.

For Theorems 2 and 3 let us denote by $\mathscr{L}_r\,(-\infty,\infty)$ the collection of functions f such that $|x|^{1-2/r}f(x) \in L_r\,(-\infty,\infty)$.

THEOREM 2. *A necessary and sufficient condition that a function* $F \in L_q\,(-\infty,\infty)$, $q > 2$, *be the Fourier transform of a function in* $\mathscr{L}_q\,(-\infty,\infty)$, $q > 2$, *is that there exist a constant* M *such that*

$$\int_{-\infty}^{\infty} |t|^{q-2}|\mathfrak{F}_{k,t}[F]|^q dt < M, \qquad k > q-2.$$

Proof of necessity. Suppose F is the Fourier transform of $f \in \mathscr{L}_q\,(-\infty,\infty)$. Then as in the proof of Theorem 1, for $t > 0$ and $k > q-2$

$$|\mathfrak{F}_{k,t}[F]| < \left\{(k/t)^{k+1}(k!)^{-1}\int_0^{\infty} e^{-ky/t}y^k|f(y)|^q dy\right\}^{1/q}$$

and consequently if $k > q-2$

$$\begin{aligned}\int_0^{\infty} t^{q-2}|\mathfrak{F}_{k,t}[F]|^q dt &< \frac{k^{k+1}}{k!}\int_0^{\infty} t^{q-k-3}dt\int_0^{\infty} e^{-ky/t}y^k|f(y)|^q dy\\ &= \frac{k^{k+1}}{k!}\int_0^{\infty} y^k|f(y)|^q dy\int_0^{\infty} e^{-ky/t}\,t^{q-k-3}dt\\ &= K(k)\int_0^{\infty} y^{q-2}\,|f(y)|^q\,dy,\end{aligned}$$

where $K(k) = k^{q-1}\,\Gamma(k-q+2)/k!$ Similarly

$$\int_{-\infty}^{0} |t|^{q-2}|\mathfrak{F}_{k,t}[F]|^q dt < K(k)\int_{-\infty}^{0} |y|^{q-2}\,|f(y)|^q dy,$$

so that

$$\int_{-\infty}^{\infty} |t|^{q-2}\,|\mathfrak{F}_{k,t}[F]|^q dt < K(k)\int_{-\infty}^{\infty} |y|^{q-2}\,|f(y)|^q dy.$$

But from Stirling's formula,

$$\lim_{k\to\infty} K(k) = 1,$$

so that $K(k)$ is bounded for $k > q-2$. Hence there is an M such that

$$\int_{-\infty}^{\infty} |t|^{q-2}\,|\mathfrak{F}_{k,t}[F]|^q dt < M, \qquad k > q-2.$$

Proof of sufficiency. Let g_+ and g_- be defined as in the proof of Theorem 1. Then as in that proof, for $t > 0$

$$L_{k,t}[g_+] = \mathfrak{F}_{k,t}[F],$$

and

$$L_{k,t}[g_-] = \mathfrak{F}_{k,-t}[F],$$

and hence

$$\int_0^\infty t^{q-2}|L_{k,t}[g_+]|^q dt = \int_0^\infty t^{q-2}|\mathfrak{F}_{k,t}[F]|^q dt < M, \qquad k > q-2$$

and

$$\int_0^\infty t^{q-2}|L_{k,t}[g_-]|^q dt = \int_0^\infty t^{q-2}|\mathfrak{F}_{k,t}[F]|^q dt < M, \qquad k > q-2.$$

Consider first g_+. By (**3**, chapter 1, Theorem 17a), with $\alpha_k(t) = t^{1-2/q}L_{k,t}[g_+]$, there is a function f_+ with $t^{1-2/q}f_+(t) \in L_q\ (0, \infty)$, and an increasing unbounded sequence of integers $\{k_i\}$ such that for any function $\beta(t) \in L_p\ (0, \infty)$,

$$\lim_{i\to\infty}\int_0^\infty \beta(t)\, t^{1-2/q}\, L_{k_i,t}[g_+]dt = \int_0^\infty \beta(t)\, t^{1-2/q} f_+(t)dt.$$

But for each $s > 0$, $t^{-(1-2/q)}\, e^{-st} \in L_p\ (0, \infty)$, and hence choosing this as our $\beta(t)$ we have for $s > 0$

$$\lim_{i\to\infty}\int_0^\infty e^{-st}\, L_{k_i,t}[g_+]dt = \int_0^\infty e^{-st} f_+(t)dt.$$

However, for $x > 0$,

$$\int_0^x |L_{k,t}[g_+]|dt < \left\{\int_0^x t^{p-2}dt\right\}^{1/p}\left\{\int_0^x t^{q-2}\, |L_{k,t}[g_+]|^q dt\right\}^{1/q}$$
$$< (p-1)^{-1/p}\, M\, x^{1/q} = O(x) \qquad \text{as} \qquad x \to \infty,$$

and as in the proof of Theorem 1, $g_+(s) \to 0$ as $s \to \infty$. Hence by (**3**, chapter 7, Theorem 11b),

$$\lim_{i\to\infty}\int_0^\infty e^{-st}\, L_{k_i,t}[g_+]dt = g_+(s), \qquad s > 0,$$

and thus

$$g_+(s) = \int_0^\infty e^{-st} f_+(t)dt, \qquad s > 0.$$

Similarly f_- exists with $t^{1-2/q}f_-(t) \in L_q\ (0, \infty)$ such that

$$g_-(s) = \int_0^\infty e^{-st} f_-(t)dt, \qquad s > 0.$$

Let

$$f(t) = \begin{cases} f_+(t), & t > 0, \\ f_-(-t), & t < 0. \end{cases}$$

Then clearly $f \in \mathscr{L}_q\ (-\infty, \infty)$, and hence by (**2**, Theorem 79) f has a Fourier transform $F^* \in L_q\ (-\infty, \infty)$. It remains to show $F = F^*$ a.e., which now follows as in Theorem 1.

THEOREM 3. *A necessary and sufficient condition that a function $F \in \mathscr{L}_p\,(-\infty, \infty)$, $1 < p < 2$, be the Fourier transform of a function in $L_p\,(-\infty, \infty)$ is that there exist a constant M such that*

$$\int_{-\infty}^\infty |\mathfrak{F}_{k,t}[F]|^p dt < M, \qquad k = 1, 2, \ldots.$$

Proof of necessity. If $F \in \mathcal{L}_p(-\infty, \infty)$ is the Fourier transform of $f \in L_p(-\infty, \infty)$ then by **(2**, Theorem 74**)**, $F \in L_q(-\infty, \infty)$, and hence by Theorem 1, there is a constant M so that

$$\int_{-\infty}^{\infty} |\mathfrak{F}_{k,t}[F]|^p dt < M, \qquad k = 1, 2, \ldots .$$

Proof of sufficiency. Let $g_+(s)$ and $g_-(s)$ be defined as in Theorem 1. Then as in that theorem,

$$\int_0^{\infty} |L_{k,t}[g_+(s)]|^p dt < M, \qquad k = 1, 2, \ldots ,$$

and

$$\int_0^{\infty} |L_{k,t}[g_-(s)]|^p dt < M, \qquad k = 1, 2, \ldots .$$

Further $g_{\pm}(s) \to 0$ as $s \to \infty$. For from Hölder's inequality we have for $s > 0$

$$|g_{\pm}(s)| < \left\{\int_{-\infty}^{\infty} \frac{|x|^{q-2}}{(x^2+s^2)^{q/2}} dx\right\}^{1/q} \left\{\int_{-\infty}^{\infty} |x|^{p-2}|F(x)|^p dx\right\}^{1/p} = 0(s^{-1/q}).$$

Hence by **(3**, chapter 7, Theorem 15a**)**, there are functions f_+ and f_- in $L_p(0, \infty)$ such that

$$g_+(s) = \int_0^{\infty} e^{-st} f_+(t) dt, \qquad s > 0$$

and

$$g_-(s) = \int_0^{\infty} e^{-st} f_-(t) dt, \qquad s > 0.$$

Let

$$f(t) = \begin{cases} f_+(t), & t > 0, \\ f_-(-t), & t < 0. \end{cases}$$

Then clearly $f \in L_p(-\infty, \infty)$ and hence by **(2**, Theorems 75 and 80**)** f has a Fourier transform $F^* \in \mathcal{L}_p(-\infty, \infty)$. It remains to show that $F = F^*$ a.e., and this follows as in Theorem 1.

References

1. E. J. McShane, *Integration* (Princeton, 1944).
2. E. C. Titchmarsh, *An introduction to the theory of Fourier integrals* (2nd ed.; Oxford, 1948).
3. D. V. Widder, *The Laplace transform* (Princeton, 1941).

University of Toronto

THE ELLIPTIC INTEGRALS OF THE THIRD KIND

E. H. NEVILLE

This paper develops a case for adopting as the standard elliptic integrals of the third kind the function $\Pi\text{s}(u, a)$ defined by

$$\Pi\text{s}(u, a) = \int_0^u \frac{\text{qs}\, a\, \text{qs}'\, a\, du}{\text{qs}^2\, u - \text{qs}^2\, a}$$

and the three functions $\Pi\text{s}\,(u, a + K_c)$, $\Pi\text{s}(u, a + K_n)$, $\Pi\text{s}(u, a + K_d)$ where K_c, K_n, K_d are the three quarter-periods of the Jacobian system. The function $\Pi\text{s}(u, a)$ is the same function whether qs u is cs u, ns u, or ds u.

The origin of the paper was a wish to understand how it has come about that the integrals commonly accepted as standard are not related symmetrically to the theta functions in terms of which they are expressed. The explanation of this irregularity is in three parts:

(1) The first of Jacobi's formulae for evaluating an elliptic integral is a deduction from the identity

$$\frac{\Theta^2 0\, \Theta(u + a)\, \Theta(u - a)}{\Theta^2 a\, \Theta^2 u} = 1 - c\, \text{sn}^2\, a\, \text{sn}^2\, u. \tag{0.1}$$

(2) To cover the range of real integrals with real variables it is necessary to use in addition to $\Theta(u + a)\, \Theta(u - a)$ the three products

$$\Theta_1(u + a)\, \Theta_1(u - a),\ \text{H}(u + a)\, \text{H}(u - a),\ \text{H}_1(u + a)\, \text{H}_1(u - a).$$

(3) If the only elliptic functions recognized are sn u, cn u, dn u, the only denominator which can be associated with the products in (2) is $\Theta^2 a\, \Theta^2 u$.

The third part of this answer is the mischief-maker leading to a set of integrals with no community of structure.

1. The notation is the systematic notation used in my *Jacobian Elliptic Functions* **(8)**, including that for bipolar functions suggested in the preface (p. iv) to the second edition (1951). Except that he prefers ω_p to K_p, it is adopted by Lenz in his paper **(7)** written as a tribute to Faber. Glaisher's function pq u is the function with simple zeros congruent with K_p and simple poles congruent with K_q and with 1 for its leading coefficient at the origin.

The bipolar function bpq u has simple poles congruent with K_p and K_q and simple zeros congruent with the other two of the four points K_s, K_c, K_n, K_d; since these other points are the zeros of the derivative pq$'$ u, the bipolar function is a multiple of the logarithmic derivative pq$'$ u/pq u and we obtain

Received July 6, 1958.

a definite function by again requiring the leading coefficient at the origin to be 1. Then

(1.1) $$\operatorname{bps} u = -\operatorname{ps}' u/\operatorname{ps} u = \operatorname{sp}' u/\operatorname{sp} u$$

and if the origin is neither pole nor zero

(1.2) $$\operatorname{bpq} u = \operatorname{sp}^2 K_q \operatorname{pq}' u/\operatorname{pq} u.$$

Explicitly, $\operatorname{bpq} u = \operatorname{rp} u \operatorname{tq} u = \operatorname{tp} u \operatorname{rq} u$, but more often than not the arbitrary coupling of a zero with a pole is an irrelevant nuisance. Since $\operatorname{ps} u \operatorname{ps}(u + K_p)$ is independent of u, (1.1) implies

(1.3) $$\operatorname{bps}(u + K_p) = -\operatorname{bps} u.$$

The theta functions I use also have 1 for leading coefficient at the origin. For $Hu/H'0$, H_1u/H_10, $\Theta u/\Theta 0$, $\Theta_1 u/\Theta_1 0$ I write $\vartheta_s u$, $\vartheta_c u$, $\vartheta_n u$, $\vartheta_d u$, relieving the memory by associating each of the functions with its lattice of zeros. The quotient $\vartheta_p u/\vartheta_q u$ is the elliptic function pq u.

The quarter-period relations between the theta functions are

(1.4, 1.5) $$\vartheta_c u = A\vartheta_s(u + K_c), \vartheta_n u = Be^{\lambda u}\vartheta_s(u + K_n)$$

(1.6) $$\vartheta_d u = C\vartheta_n(u + K_c) = De^{\lambda u}\vartheta_s(u + K_d)$$

where A, B, C, D, λ are constants whose values are not needed in this paper. From these relations it follows that the function zp u defined according to Lenz's notation **(7)** by

(1.7) $$\operatorname{zp} u = \vartheta_p' u/\vartheta_p u$$

satisfies the quarter-period relations

(1.8) $$\operatorname{zc} u = \operatorname{zs}(u + K_c), \qquad \operatorname{zd} u = \operatorname{zn}(u + K_c),$$

(1.9) $$\operatorname{zn} u = \operatorname{zs}(u + K_n) + \lambda, \qquad \operatorname{zd} u = \operatorname{zs}(u + K_d) + \lambda.$$

Since $\vartheta_n u$ is a multiple of Θu, the logarithmic derivative zn u is identical with the function Zu defined by Jacobi.

2. In terms of the function $\vartheta_n u$, Jacobi's identity (0.1) becomes

(2.1) $$\frac{\vartheta_n(a + u)\vartheta_n(a - u)}{\vartheta_n^2 a\vartheta_n^2 u} = 1 - c \operatorname{sn}^2 a \operatorname{sn}^2 u \equiv \Delta_n$$

and if we alter the numerators in turn, but not the denominator, we have

(2.2) $$\frac{\vartheta_d(a + u)\vartheta_d(a - u)}{\vartheta_n^2 a\vartheta_n^2 u} = c \operatorname{cn}^2 a \operatorname{cn}^2 u + c' \equiv \Delta_d,$$

(2.3) $$\frac{\vartheta_s(a + u)\vartheta_s(a - u)}{\vartheta_n^2 a\vartheta_n^2 u} = \operatorname{sn}^2 a - \operatorname{sn}^2 u \equiv \Delta_s,$$

(2.4) $$\frac{\vartheta_c(a + u)\vartheta_c(a - u)}{\vartheta_n^2 a\vartheta_n^2 u} = c^{-1} \operatorname{dn}^2 a \operatorname{dn}^2 u - c^{-1} c' \equiv \Delta_c.$$

It was all but inevitable that before the discovery by Glaisher in 1882 of the complete group of twelve Jacobian functions the integrands to be associated with Jacobi's integrand

$$I_n \equiv -\tfrac{1}{2}\,\partial \log \Delta_n/\partial a = c \operatorname{sn} a \operatorname{cn} a \operatorname{dn} a \operatorname{sn}^2 u/\Delta_n \tag{2.5}$$

should be

$$I_d \equiv -\tfrac{1}{2}\,\partial \log \Delta_d/\partial a = c \operatorname{sn} a \operatorname{cn} a \operatorname{dn} a \operatorname{cn}^2 u/\Delta_d \tag{2.6}$$

$$I_s \equiv -\tfrac{1}{2}\,\partial \log \Delta_s/\partial a = -\operatorname{sn} a \operatorname{cn} a \operatorname{dn} a/\Delta_s \tag{2.7}$$

$$I_c \equiv -\tfrac{1}{2}\,\partial \log \Delta_c/\partial a = \operatorname{sn} a \operatorname{cn} a \operatorname{dn} a \operatorname{dn}^2 u/\Delta_c \tag{2.8}$$

but a revision in the light of Glaisher's discovery is long overdue.

3. If

$$\Lambda_p \equiv \Lambda_p(u, a) = \tfrac{1}{2} \log \frac{\vartheta_p(a-u)}{\vartheta_p(a+u)} \tag{3.1}$$

then

$$\frac{\partial \log \Delta_p}{\partial a} = -2\frac{\partial \Lambda_p}{\partial u} - 2 \operatorname{zn} a \tag{3.2}$$

and therefore

$$\int_0^u I_p\, du = \Lambda_p(u, a) + u \operatorname{zn} a. \tag{3.3}$$

This is Jacobi's argument. The relation between the integrals is clear if we replace (3.3) by

$$\int_0^u (I_p - \operatorname{zn} a)du \equiv \Lambda_p(u, a) \tag{3.4}$$

but zn a is not an elliptic function of a, and we can only regard the integrals in (3.3) as forming not one set of peculiar interest but one of the four sets of the more general form $\Lambda_p(u, a) + u \operatorname{zq} a$.

So much was evident a century ago, and Enneper (**2**, §34) recorded the integrands corresponding to the sixteen combinations. The calculation is simple. Since $\operatorname{zq} a - \operatorname{zn} a = \operatorname{qn}'a/\operatorname{qn} a$

$$\Lambda_p(u, a) + u \operatorname{zq} a = \int_0^u \left(I_p + \frac{\operatorname{qn}'a}{\operatorname{qn} a}\right) du = -\tfrac{1}{2}\int_0^u \frac{\partial}{\partial a}\left(\log \frac{\Delta_p}{\operatorname{qn}^2 a}\right) du. \tag{3.5}$$

For given p, and q other than n, the denominator Δ_p can be put into the form $U_{pq} \operatorname{qn}^2 a + V_{pq}$, where U_{pq}, V_{pq} do not involve a, and then

$$\frac{\partial}{\partial a}\left(\log \frac{\Delta_p}{qn^2 a}\right) = -\frac{2V_{pq}\operatorname{qn}'a}{\Delta_p \operatorname{qn} a}. \tag{3.6}$$

Hence, for q other than n,

$$\Lambda_p(u, a) + u \operatorname{zq} a = \frac{\operatorname{qn}'a}{\operatorname{qn} a}\int_0^u \frac{V_{pq}\, du}{\Delta_p} \tag{3.7}$$

and the integrands which yield the sixteen integrands are given in terms of sn u, cn u, dn u compactly and explicitly in Table I.

TABLE I

DERIVATIVE OF $\frac{1}{2}\log\frac{\vartheta_p(a-u)}{\vartheta_p(a+u)}+u\frac{\vartheta_q'a}{\vartheta_q a}$ WITH RESPECT TO u

$p \backslash q$	s	c	n	d	
s	$-\mathrm{sn}^2 u$	$\mathrm{cn}^2 u$	-1	$c^{-1}\mathrm{dn}^2u$	$\div\ \mathrm{sn}^2a - \mathrm{sn}^2u$
c	cn^2u	$-c'\mathrm{sn}^2u$	dn^2u	$-c^{-1}c'$	$\div\ c^{-1}\mathrm{dn}^2a\,\mathrm{dn}^2u - c^{-1}c'$
n	1	dn^2u	$c\ \mathrm{sn}^2u$	cn^2u	$\div\ 1 - c\ \mathrm{sn}^2a\ \mathrm{sn}^2u$
d	dn^2u	c'	$c\ \mathrm{cn}^2u$	$c'\mathrm{sn}^2u$	$\div\ c\ \mathrm{cn}^2a\ \mathrm{cn}^2u + c'$
	$\times\mathrm{sn}'a/\mathrm{sn}a$	$\times\mathrm{cn}'a/\mathrm{cn}a$	$\times\mathrm{sn}a\mathrm{cn}a\mathrm{dn}a$	$\times\mathrm{dn}'a/\mathrm{dn}a$	

As functions of u, the integrands in this table are multiples of the sixteen fractions each of which has one of the four numerators 1, sn^2u, cn^2u, dn^2u and one of the four denominators Δ_s, Δ_c, Δ_n, Δ_d. In this sense the set is complete, but the structure, so clear from the integrals, is utterly obscure when only the integrands are displayed.

4. In using (3.7) we have completed our table of integrands from its third column, but since zq u − zr u = qr′ u/qr u, we could as easily complete a row from any one of its members, and we now ask if a different choice of standard integrals and a free use of Glaisher's notation will clarify the pattern of the integrands.

The clue is in the effect of quarter-period addition on the theta functions. A quarter-period addition to a is a quarter-period addition to the arguments of the two theta functions in Λ_p and to the arguments of the two theta functions in zq a, and if p and q are the same, only one transformation is involved. Let us then define a set of four integrals by writing

$$\Pi\mathrm{p}(u, a) = \Lambda_p(u, a) + u\ \mathrm{zp}\ a \tag{4.1}$$

and complete the set of integrals by means of the identity

$$\Lambda_p(u, a) + u\ \mathrm{zq}\ a = \Pi\mathrm{p}(u, a) + u\ \mathrm{qp}'a/\mathrm{qp}\ a, \tag{4.2}$$

From (1.4) and (1.6) applied to (3.1)

$$\Lambda_c(u, a) = \Lambda_s(u, a + K_c), \qquad \Lambda_d(u, a) = \Lambda_n(u, a + K_c)$$

and therefore from (1.8)

$$\Pi\mathrm{c}(u, a) = \Pi\mathrm{s}(u, a + K_c), \qquad \Pi\mathrm{d}(u, a) = \Pi\mathrm{n}(u, a + K_c). \tag{4.3, 4.4}$$

Also from (1.5)

$$\frac{\vartheta_n(a-u)}{\vartheta_n(a+u)} = e^{-2\lambda u}\frac{\vartheta_s(a+K_n-u)}{\vartheta_s(a+K_n+u)}, \frac{\vartheta_n'a}{\vartheta_n a} = \lambda + \frac{\vartheta_s'(a+K_n)}{\vartheta_s\,(a+K_n)},$$

and therefore

$$\Lambda_n(u, a) = -\lambda u + \Lambda_s(u, a + K_n), \ \mathrm{zn}\, a = \lambda + \mathrm{zs}(a + K_n),$$

implying

$$\Pi\mathrm{n}(u, a) = \Pi\mathrm{s}(u, a + K_n). \tag{4.5}$$

As functions of a, the integrands with which we are dealing are periodic in $2K_c$ and $2K_n$; hence

$$\Pi\mathrm{s}(u, a + K_c + K_n) = \Pi\mathrm{s}(u, a + K_d),$$

and from (4.4) and (4.5)

$$\Pi\mathrm{d}(u, a) = \Pi\mathrm{s}(u, a + K_d). \tag{4.6}$$

Thus for $p = c, n, d$,

$$\Pi\mathrm{p}(u, a) = \Pi\mathrm{s}(u, a + K_p). \tag{4.7}$$

In my book, $\Pi\mathrm{p}(u, a)$ is defined by this formula, and not directly in terms of the theta function $\vartheta_p u$.

The structure of the set of integrals

$$\Pi\mathrm{s}(u, a),\ \Pi\mathrm{c}(u, a),\ \Pi\mathrm{n}(u, a),\ \Pi\mathrm{d}(u, a)$$

is symmetrical, for if $\Pi\mathrm{p}(u, a)$ is any one of the four functions, then

$$\Pi\mathrm{p}(u, a),\ \Pi\mathrm{p}(u, a + K_c),\ \Pi\mathrm{p}(u, a + K_n),\ \Pi_p(u, a + K_d)$$

are the same four functions looked at, so to speak, from K_p. To put the matter differently, the symmetrical relation

$$\Pi\mathrm{q}(u, a + K_p) = \Pi\mathrm{p}(u, a + K_q) \tag{4.8}$$

shows that no one of the functions dominates the set. Briot and Bouquet (**1**, p. 447) complete the set from $\Pi\mathrm{n}(u, a)$ and associate each function $\Pi\mathrm{n}(u, a + K_q)$ with one theta function and each difference $\Pi\mathrm{n}(u, a + K_q) - \Pi\mathrm{n}(u, a)$ with one elliptic function, but their notation does not achieve the economy of typical formulae.

5. To use an integral we must be able to recognize the integrand. We denote the integrand corresponding to $\Pi\mathrm{p}(u, a)$ by J_p or if necessary by $J_p(u, a)$. In terms of theta functions

$$J_p = \partial\Lambda_p/\partial u + \mathrm{zp}\, a,$$

but what we have to consider is the explicit expression of J_p as an elliptic function. The four integrands satisfy the same quarter-period relations as the functions from which they are derived or, in other words, satisfy the typical relation

$$J_q(u, a + K_p) = J_p(u, a + K_q) \tag{5.1}$$

derived from (4.8).

In our table in §3, the functions J_s, J_c, J_n, J_d occupy the principal diagonal where they appear as follows:

$$(5.21, 5.22)\quad J_s = \frac{\mathrm{sn}'a\,\mathrm{sn}^2u}{\mathrm{sn}\,a(\mathrm{sn}^2a - \mathrm{sn}^2u)},\quad J_c = -\frac{c'\,\mathrm{cn}'a\,\mathrm{sn}^2u}{\mathrm{cn}\,a(c^{-1}\mathrm{dn}^2a\,\mathrm{dn}^2u - c^{-1}c')},$$

$$(5.23, 5.24)\quad J_n = \frac{c\,\mathrm{sn}\,a\,\mathrm{cn}\,a\,\mathrm{dn}\,a\,\mathrm{sn}^2u}{1 - c\,\mathrm{sn}^2a\,\mathrm{sn}^2u},\quad J_d = \frac{c'\,\mathrm{dn}'a\,\mathrm{sn}^2u}{\mathrm{dn}\,a(c\,\mathrm{cn}^2a\,\mathrm{cn}^2u + c')}.$$

We may suggest that it is because only the original Jacobian functions sn u, cn u, dn u are used that the symmetry of the quartette cannot be seen, but since each function might be expressed in terms of any one of the twelve functions pq u, we are not likely to find satisfactory transformations by a process of trial and error.

We take a hint from the Weierstrassian theory, in which the fundamental integrand of the third kind is $\wp'a/(\wp u - \wp a)$, and we have

$$\int_0^u \frac{\wp'a\,du}{\wp u - \wp a} = \log\frac{\sigma(a-u)}{\sigma(a+u)} + \frac{2u\,\sigma'a}{\sigma a}.$$

If the Weierstrassian functions have the same lattice as the Jacobian functions, $\vartheta_s u$ and σu are integral functions with the same zeros, and the relation between them is

$$\sigma u = e^{\mu u^2}\vartheta_s u,$$

where μ is a constant. Hence

$$\log\frac{\sigma(a-u)}{\sigma(a+u)} = -4\,\mu a u + \log\frac{\vartheta_s(a-u)}{\vartheta_s(a+u)},\quad \frac{\sigma'a}{\sigma\,a} = 2\,\mu a + \frac{\vartheta_s'a}{\vartheta_s\,a},$$

and therefore

$$\int_0^u \frac{\wp'a\,du}{\wp u - \wp a} = 2\{\Lambda_s(u, a) + u\,\mathrm{zs}\,a\},$$

that is,

$$J_s = \frac{\frac{1}{2}\wp'a}{\wp u - \wp a}.$$

Since $\wp u$ differs from qs^2u by a constant, whether q is c, n, or d, we have

$$(5.3)\qquad J_s = \frac{\mathrm{qs}\,a\,\mathrm{qs}'a}{\mathrm{qs}^2u - \mathrm{qs}^2a}$$

and therefore

$$(5.4)\qquad J_p(u, a) = \frac{\mathrm{qs}(a + K_p)\,\mathrm{qs}'(a + K_p)}{\mathrm{qs}^2u - \mathrm{qs}^2(a + K_p)},$$

a general formula which includes (5.3).

To verify that the formulae (5.21—5.24) extracted from the table in §3 can be deduced from (5.4) is an exercise in algebra. First, $\mathrm{qs}'a = -\mathrm{sq}'a/\mathrm{sq}^2a$, gives

$$(5.5)\qquad J_s = \frac{\mathrm{sq}'a\,\mathrm{sq}^2u}{\mathrm{sq}\,a(\mathrm{sq}^2a - \mathrm{sq}^2u)};$$

this formula includes (5.21), and shows that in spite of appearances the integrand given by (5.21) does not stand in any special relation to K_n.

Next, since $\mathrm{ps}\,a(\mathrm{ps}\,a + K_p) = \mathrm{ps}'K_p$, identification of q with p in (5.4) gives

$$J_p = \frac{\mathrm{ps}'^2K_p\,\mathrm{sp}\,a\,\mathrm{sp}'a}{\mathrm{ps}^2u - \mathrm{ps}'^2K_p\,\mathrm{sp}^2a},$$

that is,

$$J_p = \frac{\mathrm{ps}'^2K_p\,\mathrm{sp}\,a\,\mathrm{sp}'a\,\mathrm{sp}^2u}{1 - \mathrm{ps}'^2K_p\,\mathrm{sp}^2a\,\mathrm{sp}^2u}; \tag{5.61}$$

this formula includes (5.23), identifying Jacobi's integrand with $J_s(u, a + K_n)$. In other words, $\Pi\mathrm{n}(u, a)$ is Jacobi's function $\Pi(u, a)$ seen as one member of a set of which the other two members $\Pi\mathrm{c}(u, a)$ $\Pi\mathrm{d}(u, a)$ have their integrands given by

$$(5.62,\ 5.63)\qquad J_c = \frac{c'\,\mathrm{sc}\,a\,\mathrm{sc}'a\,\mathrm{sc}^2u}{1 - c'\,\mathrm{sc}^2a\,\mathrm{sc}^2u}, \qquad J_d = -\frac{cc'\,\mathrm{sd}\,a\,\mathrm{sd}'a\,\mathrm{sd}^2u}{1 + cc'\,\mathrm{sd}^2a\,\mathrm{sd}^2u}.$$

Lastly, to recover (5.22) and (5.24) from (5.4), we suppose q to be distinct from p and r to be the third member of the set c, n, d; then

$$(5.71,\ 5.72)\qquad \mathrm{qs}(a + K_p) = \mathrm{qs}K_p\,\mathrm{rp}\,a, \qquad \mathrm{qr}(a + K_p) = \mathrm{qr}K_p\mathrm{rq}\,a.$$

Since $\mathrm{sr}^2u\{\mathrm{qs}^2u - \mathrm{qs}^2(a + K_p)\}$ is a linear function of qr^2u which is zero only if $\mathrm{qr}^2u = \mathrm{qr}^2(a + K_p)$, it follows that $\mathrm{qs}^2u - \mathrm{qs}^2(a + K_p)$ is a multiple of $\mathrm{rs}^2u(\mathrm{qr}^2a\,\mathrm{qr}^2u - \mathrm{qr}^2K_p)$, and is therefore the product of

$$\mathrm{rs}^2u(\mathrm{pq}^2K_r\,\mathrm{qr}^2a\,\mathrm{qr}^2u + \mathrm{pr}^2K_q)$$

by a factor independent of u. Determining the factor by putting $u = K_q$ and using (5.71), we have

$$\mathrm{qs}^2u - \mathrm{qs}^2(a + K_p) = \mathrm{rp}^2a\,\mathrm{rs}^2u(\mathrm{pq}^2K_r\mathrm{qr}^2a\,\mathrm{qr}^2u + \mathrm{pr}^2K_q).$$

Using (5.71) again and replacing $\mathrm{qs}^2K_p\,\mathrm{rp}'a/\mathrm{rp}\,a$ by $\mathrm{ps}^2K_q\,\mathrm{pr}'a/\mathrm{pr}\,a$, we have

$$J_p = \frac{\mathrm{ps}^2K_q\,\mathrm{pr}'a\,\mathrm{sr}^2u}{\mathrm{pr}\,a(\mathrm{pq}^2K_r\,\mathrm{qr}^2a\,\mathrm{qr}^2u + \mathrm{pr}^2K_q)}. \tag{5.81}$$

This is the formula of which (5.22) and (5.24) are two cases; a third case is another formula for Jacobi's integrand:

$$J_n = \frac{c\,\mathrm{nc}'a\,\mathrm{sc}^2u}{\mathrm{nc}\,a(c'^{-1}\mathrm{dc}^2a\,\mathrm{dc}^2u - cc'^{-1})}.$$

In fact there are six cases of (5.81), but the interchange of q and r is almost trivial. The direct transformation of (5.82) into (5.23) takes the form

$$\frac{c\,\mathrm{nc}'a\,\mathrm{sc}^2u}{\mathrm{nc}\,a(c'^{-1}\,\mathrm{dc}^2a\,\mathrm{dc}^2u - cc'^{-1})} = -\frac{cc'\,\mathrm{cn}\,a\,\mathrm{cn}'a\,\mathrm{sn}^2u}{\mathrm{dn}^2a\,\mathrm{dn}^2u - c\,\mathrm{cn}^2a\,\mathrm{cn}^2u}$$
$$= \frac{cc'\,\mathrm{sn}\,a\,\mathrm{sn}'a\,\mathrm{sn}^2u}{(1 - c\,\mathrm{sn}^2a)(1 - c\,\mathrm{sn}^2u) - c(1 - \mathrm{sn}^2a)(1 - \mathrm{sn}^2u)}.$$

6. The relation between $\Pi p(u,a)$ and $\Pi q(u,a)$ can be expressed as a relation between functions instead of as a relation between arguments, for (4.1) gives

$$\Pi p(u,a) - \Pi q(u,a) = \tfrac{1}{2}\log\frac{\mathrm{pq}(a-u)}{\mathrm{pq}(a+u)} + u\,.\,\frac{\mathrm{pq}'a}{\mathrm{pq}\,a}. \tag{6.1}$$

In other words, an alternative definition of $\Pi p(u,a)$ in terms of $\Pi s(u,a)$ is

$$\Pi p(u,a) = \Pi s(u,a) + \tfrac{1}{2}\log\frac{\mathrm{ps}(a-u)}{\mathrm{ps}(a+u)} + u\,.\,\frac{\mathrm{ps}'a}{\mathrm{ps}\,a}. \tag{6.2}$$

The additional logarithmic ambiguity is only apparent if it is understood that the logarithm is zero when $u = 0$ and varies continuously as u describes the path of integration implicit in $\Pi s(u,a)$.

It is interesting to establish (6.2) in terms of integrands. With differences of notation, the algebra is essentially Legendre's (**5**, §46; **6**, §49). With the use of the bipolar function, the addition theorem for ps u can be written

$$\mathrm{ps}(u+v) = \mathrm{ps}\,u\,\mathrm{ps}\,v(\mathrm{bps}\,u - \mathrm{bps}\,v)/(\mathrm{ps}^2u - \mathrm{ps}^2v).$$

Hence

$$\frac{\mathrm{ps}(a-u)}{\mathrm{ps}(a+u)} = \frac{\mathrm{bps}\,u + \mathrm{bps}\,a}{\mathrm{bps}\,u - \mathrm{bps}\,a}, \tag{6.3}$$

and the result to be proved is, that if $a_p = a + K_p$, then

$$\frac{\mathrm{ps}\,a_p\,\mathrm{ps}'a_p}{\mathrm{ps}^2u - \mathrm{ps}^2a_p} = \frac{\mathrm{ps}\,a\,\mathrm{ps}'a}{\mathrm{ps}^2u - \mathrm{ps}^2a} - \frac{\mathrm{bps}\,a\,\mathrm{bps}'u}{\mathrm{bps}^2u - \mathrm{bps}^2a} + \frac{\mathrm{ps}'a}{\mathrm{ps}\,a};$$

since

$$\mathrm{ps}'a = -\mathrm{ps}\,a\,\mathrm{bps}\,a,\ \mathrm{ps}'a_p = -\mathrm{ps}\,a_p\,\mathrm{bps}\,a_p = \mathrm{ps}\,a_p\,\mathrm{bps}\,a.$$

From (1.3), this is equivalent to

$$-\frac{\mathrm{bps}'u}{\mathrm{bps}^2u - \mathrm{bps}^2a} = 1 + \frac{\mathrm{ps}^2a}{\mathrm{ps}^2u - \mathrm{ps}^2a} + \frac{\mathrm{ps}^2a_p}{\mathrm{ps}^2u - \mathrm{ps}^2a_p}. \tag{6.4}$$

Now $\mathrm{ps}^2u(\mathrm{bps}^2u - \mathrm{bps}^2a)$ is a quadratic function of ps^2u which vanishes if $\mathrm{ps}^2u = \mathrm{ps}^2a$ and therefore also, from (1.3), if $\mathrm{ps}^2u = \mathrm{ps}^2a_p$; also the coefficient of ps^4u in $\mathrm{ps}\,u\,\mathrm{bps}^2u$, that is, in ps'^2u, is 1. Hence

$$\mathrm{ps}^2u(\mathrm{bps}^2u - \mathrm{bps}^2a) = (\mathrm{ps}^2u - \mathrm{ps}^2a)(\mathrm{ps}^2u - \mathrm{ps}^2a_p). \tag{6.5}$$

Multiplying by sp^2u, differentiating, and substituting for $\mathrm{ps}'u$ and $\mathrm{sp}'u$ from (1.1), we have

$$\mathrm{bps}\,u\,\mathrm{bps}'u = -\mathrm{bps}\,u(\mathrm{ps}^2u - \mathrm{ps}^2a\,\mathrm{ps}^2a_p\,\mathrm{sp}^2u),$$

that is,

$$-\mathrm{ps}^2u\,\mathrm{bps}'u = \mathrm{ps}^4u - \mathrm{ps}^2a\,\mathrm{ps}^2a_p. \tag{6.6}$$

From (6.5) and (6.6),

$$-\frac{\mathrm{bps}'u}{\mathrm{bps}^2u - \mathrm{bps}^2a} = \frac{\mathrm{ps}^4u - \mathrm{ps}^2a\,\mathrm{ps}^2a_p}{(\mathrm{ps}^2u - \mathrm{ps}^2a)(\mathrm{ps}^2u - \mathrm{ps}^2a_p)}, \tag{6.7}$$

and the right-hand side of (6.7), resolved into partial fractions in the variable $\mathrm{ps}^2 u$, is the right-hand side of (6.4).

7. Since $\Pi\mathrm{s}(u, a)$ is an odd function of a, (6.2) can be written

$$(7.11)\qquad \Pi\mathrm{s}(u, a) + \Pi\mathrm{s}(u, K_p - a) = \tfrac{1}{2}\log\frac{\mathrm{ps}(a+u)}{\mathrm{ps}(a-u)} - u\,.\,\frac{\mathrm{ps}'a}{\mathrm{ps}\,a}\,;$$

further, since $\Pi\mathrm{s}(u, a)$ as a function of a, has $2K_q$ for a period,

$$\Pi\mathrm{s}(u, K_p - (a + K_q)) = \Pi\mathrm{s}(u, (K_p - a) + K_q),$$

and substituting $a + K_q$ for a in (7.11) we have for $\mathrm{q} \neq \mathrm{p}$,

$$(7.12)\qquad \Pi\mathrm{q}(u, a) + \Pi\mathrm{q}(u, K_p - a) = \tfrac{1}{2}\log\frac{\mathrm{rq}(a+u)}{\mathrm{rq}(a-u)} - u\,.\,\frac{\mathrm{rq}'a}{\mathrm{rq}\,a}\,.$$

The formulae (7.11) and (7.12) may be regarded as halving the area of values of a throughout which $\Pi\mathrm{q}(u, a)$ requires a theta function for its evaluation.

From these formulae we see also that if $2a$ is a quarter-period the integrals of the third kind degenerate. Since the value of $\mathrm{ps}(\tfrac{1}{2}K_p + u)\mathrm{ps}(\tfrac{1}{2}K_p - u)$ is $\mathrm{ps}^2\tfrac{1}{2}K_p$, we have from (7.11)

$$(7.21)\qquad \Pi\mathrm{s}(u, \tfrac{1}{2}K_p) = \tfrac{1}{2}\log\{\mathrm{sp}\,\tfrac{1}{2}K_p\,\mathrm{ps}(u + \tfrac{1}{2}K_p)\} + \tfrac{1}{2}\,u\,\mathrm{bps}\,\tfrac{1}{2}K_p.$$

Also $\mathrm{rq}(\tfrac{1}{2}K_p + u)\mathrm{rq}(\tfrac{1}{2}K_p - u)$ is a constant, since addition of K_p to u interchanges the poles and the zeros of $\mathrm{rq}\,u$; this constant is $\mathrm{rq}^2\,\tfrac{1}{2}K_p$, and we have from (7.12)

$$(7.22)\qquad \Pi\mathrm{q}(u, \tfrac{1}{2}K_p) = \tfrac{1}{2}\log\{\mathrm{qr}\,\tfrac{1}{2}K_p\,\mathrm{rq}(u + \tfrac{1}{2}K_p)\} - \tfrac{1}{2}u(\mathrm{bqs}\tfrac{1}{2}K_p - \mathrm{brs}\tfrac{1}{2}K_p),$$

since $\mathrm{rq}\,a = \mathrm{rs}\,a/\mathrm{qs}\,a$.

For the sake of completeness we must add that the identities

$$\Pi\mathrm{p}(u, K_p - a) = -\,\Pi\mathrm{s}(u, a),\ \Pi\mathrm{p}(u, a) = -\,\Pi\mathrm{s}(u, K_p - a)$$

imply

$$(7.31)\qquad \Pi\mathrm{p}(u, a) + \Pi\mathrm{p}(u, K_p - a) = \tfrac{1}{2}\log\frac{\mathrm{sp}(a+u)}{\mathrm{sp}(a-u)} - u\,.\,\frac{\mathrm{sp}'a}{\mathrm{sp}\,a}\,,$$

$$(7.32)\qquad \Pi\mathrm{p}(u, \tfrac{1}{2}K_p) = \tfrac{1}{2}\log\{\mathrm{ps}\,\tfrac{1}{2}K_p\,\mathrm{sp}(u + \tfrac{1}{2}K_p)\} - \tfrac{1}{2}\,u\,\mathrm{bps}\,\tfrac{1}{2}K_p.$$

To us, (7.31) and (7.32) are little more than repetitions of (7.11) and (7.21), but we must remember that since the function we are denoting by $\Pi\mathrm{n}(u, a)$ was known long before $\Pi\mathrm{s}(u, a)$ was introduced, the classical formulae implicit in Jacobi's *theorema de additione argumenti parametri* (**4**, p. 159) are cases of (7.22) and (7.32).

The values of the bipolar functions used in (7.21) and (7.22) are easily found. For any value of u,

$$(7.41)\qquad \mathrm{qs}\,2u + \mathrm{rs}\,2u = \mathrm{bps}\,u,$$

and therefore

$$(7.42,\ 7.43)\qquad \mathrm{bps}\,\tfrac{1}{2}K_p = \mathrm{qs}K_p + \mathrm{rs}K_p,\ \mathrm{bqs}\,\tfrac{1}{2}K_p = \mathrm{rs}K_p.$$

Thus (7.22) becomes

$$(7.44)\quad \Pi\mathrm{q}(u, \tfrac{1}{2}K_p) = \tfrac{1}{2}\log\{\mathrm{qr}\tfrac{1}{2}K_p\,\mathrm{rq}(u + \tfrac{1}{2}K_p)\} + \tfrac{1}{2}u(\mathrm{qs}\,K_p - \mathrm{rs}\,K_p).$$

We can modify the logarithmic terms in (7.21) and (7.22) and take fuller advantage of (7.41) and (7.42). From (6.3), (7.11) is equivalent to

$$(7.51)\quad \Pi\mathrm{s}(u, a) + \Pi\mathrm{s}(u, K_p - a) = \tfrac{1}{2}\log\frac{\mathrm{bps}\,u - \mathrm{bps}\,a}{\mathrm{bps}\,u + \mathrm{bps}\,a} + u\,\mathrm{bps}\,a,$$

and therefore (7.21) is equivalent to

$$(7.52)\quad \Pi\mathrm{s}(u, \tfrac{1}{2}K_p) = \tfrac{1}{4}\log\frac{\mathrm{bps}\,u - \mathrm{qs}K_p - \mathrm{rs}K_p}{\mathrm{bps}\,u + \mathrm{qs}K_p + \mathrm{rs}K_p} + \tfrac{1}{2}u(\mathrm{qs}K_p + \mathrm{rs}K_p).$$

Instead of (7.12) we have

$$(7.53)\quad \Pi\mathrm{q}(u, a) + \Pi\mathrm{q}(u, K_p - a)$$

$$= \tfrac{1}{2}\log\frac{(\mathrm{bqs}\,u + \mathrm{bqs}\,a)(\mathrm{brs}\,u - \mathrm{brs}\,a)}{(\mathrm{bqs}\,u - \mathrm{bqs}\,a)(\mathrm{brs}\,u + \mathrm{brs}\,a)} - u(\mathrm{bqs}\,a - \mathrm{brs}\,a),$$

leading to

$$(7.54)\quad \Pi\mathrm{q}(u, \tfrac{1}{2}K_p)$$

$$= \tfrac{1}{4}\log\frac{(\mathrm{bqs}\,u + \mathrm{rs}K_p)(\mathrm{brs}\,u - \mathrm{qs}K_p)}{(\mathrm{bqs}\,u - \mathrm{rs}K_p)(\mathrm{brs}\,u + \mathrm{qs}K_p)} + \tfrac{1}{2}u(\mathrm{qs}K_p - \mathrm{rs}K_p).$$

The squares of the constants $\mathrm{qs}K_p$ are given by

$$(7.61)\quad \mathrm{ns}^2K_c = -\mathrm{cs}^2K_n = 1,\ \mathrm{ns}^2K_d = -\mathrm{ds}^2K_n = c,\ \mathrm{ds}^2K_c = -\mathrm{cs}^2K_d = c',$$

and depend only on the Jacobian system, but the constants themselves with the exception of $\mathrm{ns}K_c$ depend on the choice of a basis for the lattice. Defining v, k, k' by

$$(7.62)\quad v = \mathrm{sc}K_n,\ k = \mathrm{ns}(K_c + K_n),\ k' = \mathrm{ds}K_c,$$

we have

$$(7.63)\quad v^2 = -1,\ k^2 = c,\ k'^2 = c',$$

and the six critical constants are given by

$$(7.64)\quad \mathrm{ns}K_c = 1,\ \mathrm{cs}K_n = -v,\ \mathrm{ns}K_d = -k,$$
$$\mathrm{ds}K_n = -vk,\ \mathrm{ds}K_c = k',\ \mathrm{cs}K_d = vk'.$$

The relations

$$\mathrm{ns}K_c/\mathrm{cs}K_n = \mathrm{ds}K_n/\mathrm{ns}K_d = \mathrm{cs}K_d/\mathrm{ds}K_c = v$$

express that rotation in the direction $K_c \to K_n \to K_d$ is positive or negative according as v is $+i$ or $-i$.

The results of expressing the constants in (7.52) and (7.54) in terms of v, k, k' are valid for all Jacobian systems, but it is for the classical systems in which k and k' are real that they are specially required.

8. In proposing that the typical integrand in the table in §3 should be treated as $J_p(u, a) + \mathrm{qp}'a/\mathrm{qp}\, a$ rather than as $I_p(u, a) + \mathrm{qn}'a/\mathrm{qn}\, a$, we are not altering the composition of the table. The integrands are the same sixteen functions of u and a, and the most to be claimed is that with the whole set of Glaisher's functions at our service we have shown that we can move easily from one entry to another within the table. To Hermite **(3)** are due examples of a process by which the tale of recorded integrals of the third kind can be quadrupled in length. The denominator Δ_n in (2.1) is the denominator in the classical expression for $\mathrm{sn}(a + u)$, and since

$$\vartheta_s(a+u)/\vartheta_n(a+u) = \mathrm{sn}(a+u) = (\mathrm{sn}\, a\, \mathrm{cn}\, u\, \mathrm{dn}\, u + \mathrm{cn}\, a\, \mathrm{dn}\, a\, \mathrm{sn}\, u)/\Delta_n$$

we have

$$\frac{\vartheta_s(a+u)\vartheta_n(a-u)}{\vartheta_n^2 a\, \vartheta_n^2 u} = \mathrm{sn}\, a\, \mathrm{cn}\, u\, \mathrm{dn}\, u + \mathrm{cn}\, a\, \mathrm{dn}\, a\, \mathrm{sn}\, u. \tag{8.1}$$

Jacobi's argument now gives

$$\int_0^u \frac{\mathrm{cn}\, a\, \mathrm{dn}\, a\, \mathrm{cn}\, u\, \mathrm{dn}\, u - \mathrm{sn}\, a(\mathrm{dn}^2 a + c\,\mathrm{cn}^2 a)\mathrm{sn}\, u}{\mathrm{sn}\, a\, \mathrm{cn}\, u\, \mathrm{dn}\, u + \mathrm{cn}\, a\, \mathrm{dn}\, a\, \mathrm{sn}\, u}\, du \tag{8.2}$$

$$= \log\frac{\vartheta_s(a+u)}{\vartheta_n(a-u)} - 2u\,.\,\frac{\vartheta_n' a}{\vartheta_n\, a}.$$

This method gives integrands corresponding to the 48 integrals

$$\tfrac{1}{2}\log\frac{\vartheta_p(a-u)}{\vartheta_r(a+u)} + u\,.\,\frac{\vartheta_q' a}{\vartheta_q\, a}$$

with $p \neq r$, but Hermite himself attached no importance to the extension. His comment, "au fond, ces diverses expressions se ramènent à la quantité . . ." $\Pi(u, a)$, suggests only that he was dissatisfied with the incoherent mass of formulae derived from Jacobi's integrand and its three companions.

More interesting than this extension is Hermite's use of the integrand

$$\mathrm{sn}\, a\, \mathrm{cn}\, a\, \mathrm{dn}\, a/(\mathrm{sn}^2 u - \mathrm{sn}^2 a),$$

which is the integrand denoted above (§§2–3) by I_s, in preference to Jacobi's integrand I_n, or, in other words, his use of the integral $\Lambda_s(u, a) + u\, \mathrm{zn}\, a$ in preference to Jacobi's integral $\Pi(u, a)$ which is $\Lambda_n(u, a) + u\, \mathrm{zn}\, a$. "Cette intégrale présente," he says, "plus de facilité que celle de Jacobi pour établir les théorèmes sur l'addition des arguments" **(3**, p. 841**)**. That is to say, he has found that the advantages of using the function $\Lambda_s(u, a)$ associated with the origin instead of the corresponding function $\Lambda_n(u, a)$ associated with the point K_n outweigh any disadvantages due to the heterogeneity of $\Lambda_s(u, a) + u\, \mathrm{zn}\, a$ as compared with $\Lambda_n(u, a) + u\, \mathrm{zn}\, a$. And this in spite of the fact that for elliptic functions he has only those whose poles are congruent with K_n.

9. The integrands tabulated in §3 are functions to which Jacobi's method of integration is seen in advance to be applicable; we have still to consider the arbitrary integrand $\lambda/(\mathrm{pq}^2 u - \mu)$. Determining a constant a by the condition

$$\mathrm{pq}^2\, a = \mu, \tag{9.11}$$

and inserting a numerator found to be convenient, we deal with the integral

$$\int_0^u \frac{\mathrm{pq}\, a\, \mathrm{pq}'a\, du}{\mathrm{pq}^2 u - \mathrm{pq}^2 a}.$$

If q is s, the integral is already known, for (5.3) is equivalent to

$$\int_0^u \frac{\mathrm{ps}\, a\, \mathrm{ps}'a\, du}{\mathrm{ps}^2 u - \mathrm{ps}^2 a} = \Pi\mathrm{s}(u, a). \tag{9.12}$$

If q is not s, then $\mathrm{pq}^2 u$ is a linear function of $\mathrm{sq}^2\, u$; whether $\mathrm{pq}^2 u$ is $\mathrm{sq}^2 u$ or $1 - \mathrm{qs}^2 K_p \mathrm{sq}^2 u$

$$\frac{\mathrm{pq}\, a\, \mathrm{pq}'a}{\mathrm{pq}^2 u - \mathrm{pq}^2 a} = \frac{\mathrm{sq}\, a\, \mathrm{sq}'a}{\mathrm{sq}^2 u - \mathrm{sq}^2 a},$$

and since

$$\frac{\mathrm{qs}\, a\, \mathrm{qs}'a}{\mathrm{qs}^2 u - \mathrm{qs}^2 a} = \frac{\mathrm{qs}'a}{\mathrm{qs}\, a} \cdot \frac{\mathrm{qs}^2 a}{\mathrm{qs}^2 u - \mathrm{qs}^2 a} = \frac{\mathrm{sq}'a}{\mathrm{sq}\, a} \cdot \frac{\mathrm{sq}^2 u}{\mathrm{sq}^2 u - \mathrm{sq}^2 a}$$

we have

$$\frac{\mathrm{pq}\, a\, \mathrm{pq}'a}{\mathrm{sq}^2 a} \int_0^u \frac{\mathrm{sq}^2 u\, du}{\mathrm{pq}^2 u - \mathrm{pq}^2 a} = \Pi\mathrm{s}(u, a), \tag{9.13}$$

$$\int_0^u \frac{\mathrm{pq}\, a\, \mathrm{pq}'a\, du}{\mathrm{pq}^2 u - \mathrm{pq}^2 a} = \frac{u\, \mathrm{qs}'a}{\mathrm{qs}\, a} + \Pi\mathrm{s}(u, a). \tag{9.14}$$

Although (9.13) is valid whether or not p is s, it is worth while to separate the two cases for the sake of further simplification. If p is s, the formula is

$$\frac{\mathrm{sq}'a}{\mathrm{sq}\, a} \int_0^u \frac{\mathrm{sq}^2 u\, du}{\mathrm{sq}^2 u - \mathrm{sq}^2 a} = \Pi\mathrm{s}(u, a), \tag{9.15}$$

a simple variant of (9.12), and if p is not s it can be written

$$\frac{\mathrm{ps}^2 K_q\, \mathrm{sq}'a}{\mathrm{sq}\, a} \int_0^u \frac{\mathrm{sq}^2 u\, du}{\mathrm{pq}^2 u - \mathrm{pq}^2 a} = \Pi\mathrm{s}(u, a), \tag{9.16}$$

since

$$\mathrm{qs}^2 K_p = -\mathrm{ps}^2 K_q.$$

The earliest of all integrals of the third kind, Legendre's function Π defined by (**5**, p. 17; **6**, p. 17)

$$\Pi = \int_0^\phi \frac{du}{(1 + n \sin^2\phi)\Delta},$$

where $\Delta = \sqrt{(1 - c\sin^2\phi)}$, is the integral

$$\int_0^u \frac{du}{1 + n\,\mathrm{sn}^2 u}.$$

It is usual now to change the sign in the denominator, and we take the integral of this form with sn u replaced by pq u as

$$\int_0^u \frac{du}{1 - \lambda\,\mathrm{pq}^2 u}.$$

If we define a by

$$\mathrm{qp}^2 a = \lambda, \tag{9.21}$$

we have

$$\frac{\mathrm{qp}'a/\mathrm{qp}\,a}{1 - \lambda\,\mathrm{pq}^2 u} = \frac{\mathrm{pq}\,a\,\mathrm{pq}'a}{\mathrm{pq}^2 u - \mathrm{pq}^2 a},$$

and we have merely to rewrite (9.12), (9.14), (9.15), and (9.16) as

$$\frac{\mathrm{sp}'a}{\mathrm{sp}\,a}\int_0^u \frac{du}{1 - \mathrm{sp}^2 a\ \mathrm{ps}^2 u} = \Pi\mathrm{s}(u, a), \tag{9.22}$$

$$\frac{\mathrm{qp}'a}{\mathrm{qp}\,a}\int_0^u \frac{du}{1 - \mathrm{qp}^2 a\,\mathrm{pq}^2 u} = \frac{u\,\mathrm{qs}'a}{\mathrm{qs}\,a} + \Pi\mathrm{s}(u, a), \tag{9.23}$$

$$\int_0^u \frac{\mathrm{qs}\,a\,\mathrm{qs}'a\,\mathrm{sq}^2 u\,du}{1 - \mathrm{qs}^2 a\,\mathrm{sq}^2 u} = \Pi\mathrm{s}(u, a), \tag{9.24}$$

$$\frac{\mathrm{ps}^2 K_q\,\mathrm{ps}'a}{\mathrm{ps}\,a}\int_0^u \frac{\mathrm{sq}^2 u\,du}{1 - \mathrm{qp}^2 a\,\mathrm{pq}^2 u} = \Pi\mathrm{s}(u, a). \tag{9.25}$$

There is an alternative substitution. The function pq u has one of the quarter-periods of the Jacobian system for a half-period, and if this quarter-period is K_t, the product qp$u$ qp$(u + K_t)$ is independent of u, that is, is a constant of the system. If the square of this constant is j_{pq}, to write

$$\lambda = j_{pq}\,\mathrm{pq}^2 a \tag{9.31}$$

is equivalent to writing

$$\mathrm{qp}^2(a + K_t) = \lambda,$$

and this change replaces $\Pi\mathrm{s}(u, a)$ by $\Pi\mathrm{t}(u, a)$. The quarter-period relevant for ps u and sp u is K_p, and if the three quarter-periods of the system are K_p, K_q, K_r, then $\mathrm{sp}(K_p + K_q) = -\mathrm{sp}K_r$ and

$$j_{ps} = \mathrm{sp}^2K_q\,\mathrm{sp}^2K_r; \qquad j_{sq} = \mathrm{qs}^2K_p\,\mathrm{qs}^2K_r.$$

The quarter-period relevant for pq u is K_r, and

$$j_{pq} = \mathrm{qp}^2K_r.$$

From (9.22), (9.24), and (9.25) we have

(9.32) $$\frac{\mathrm{ps}'a}{\mathrm{ps}\,a}\int_0^u \frac{du}{1 - j_{ps}\,\mathrm{ps}^2a\,\mathrm{ps}^2u} = \Pi\mathrm{p}(u, a),$$

(9.33) $$\int_0^u j_{sq}\frac{\mathrm{sq}\,a\,\mathrm{sq}'a\,\mathrm{sq}^2u\,du}{1 - j_{sq}\,\mathrm{sq}^2a\,\mathrm{sq}^2u} = \Pi\mathrm{q}(u, a),$$

(9.34) $$\frac{\mathrm{ps}^2K_q\,\mathrm{qr}'a}{\mathrm{qr}\,a}\int_0^u \frac{\mathrm{sq}^2u\,du}{1 - j_{pq}\,\mathrm{pq}^2a\,\mathrm{pq}^2u} = \Pi\mathrm{r}(u, a),$$

and from (9.23)

(9.35) $$\frac{\mathrm{sq}'a}{\mathrm{sq}\,a}\int_0^u \frac{du}{1 - j_{pq}\,\mathrm{pq}^2a\,\mathrm{pq}^2u} = \frac{u\,\mathrm{pr}'a}{\mathrm{pr}\,a} + \Pi\mathrm{r}(u, a).$$

We can now see the structure of the integrands which compose the leading diagonal of the table in §3. Since the only functions to be used are Jacobi's three functions sn u, cn u, dn u, the denominator has one of the two forms $\mathrm{pn}^2u - \mathrm{pn}^2a$, $1 - j_{pn}\,\mathrm{pn}^2a\mathrm{pn}^2u$. The integrand J_s corresponding to $\Pi\mathrm{s}(u, a)$ is the integrand in (9.15) with n for q; to use (9.16) would be merely to substitute $-(\mathrm{cn}^2u - \mathrm{cn}^2a)$ or $-(\mathrm{dn}^2u - \mathrm{dn}^2a)/c$ for $\mathrm{sn}^2u - \mathrm{sn}^2a$. The function $\Pi\mathrm{n}(u, a)$ comes only from (9.33), and since $j_{sn} = \mathrm{ns}^2K_c\mathrm{ns}^2(K_c + K_n) = c$, we find the integrand J_n as $c\,\mathrm{sn}\,a\,\mathrm{sn}'a\,\mathrm{sn}^2u/(1 - c\,\mathrm{sn}^2a\,\mathrm{sn}^2u)$, precisely as given by Jacobi. The functions $\Pi\mathrm{c}(u, a)$ and $\Pi\mathrm{d}(u, a)$ come from (9.34), the one when pq u is dn u and the other when pq u is cn u, but we must express $\mathrm{qr}'a/\mathrm{qr}\,a$ as $-\mathrm{rq}'a/\mathrm{rq}\,a$; the constants required are given by

$$\mathrm{ds}^2K_n = -c,\ j_{dn} = 1/c';\ \mathrm{cs}^2K_n = -1,\ j_{cn} = -c/c'$$

and the entries in the table can be verified immediately.

10. As we have said, the substitution $\mu = \mathrm{pq}^2a$ does not impose any restrictions on μ, and theoretically the two formulae (9.12) and (9.14), together with the expression of $\Pi\mathrm{s}(u, a)$ as $\Lambda_s(u, a) + u\,\mathrm{zs}\,a$, reduce any function of the third kind to a combination of functions each of which is a function of a single argument. But if the problem is the evaluation of a real integral by means of real variables, there are complications. A real value of μ does not necessarily give a real value of a, and if u is real and a complex, then functions of $a + u$ are functions which must be dissected before they can be evaluated.

In discussing evaluation, we assume that K_c has a real value K and K_n an imaginary value iK', and we assume also that K and K' are positive; then k and k' are positive, and v is i. The origin and the points K, $K + iK'$, iK' are the corners of a rectangle which we denote by *SCDN*. In applying general formulae it is important to remember that $K + iK'$ is $-K_d$, since in the formal theory the three quarter-periods satisfy the symmetrical relation $K_c + K_n + K_d = 0$.

The path of integration is a segment of the real axis. For the present we continue to take $u = 0$ for the lower limit; the effect of removing this restriction is considered in our concluding paragraph.

If one of the twelve functions $\mathrm{pq}^2 a$ is real, all of them are real, and therefore each of the functions $\mathrm{pq}\, a$ and each of the derivatives $\mathrm{pq}'a$ is either real or imaginary. Hence in all that follows each of the functions $\Pi\mathrm{p}(u, a)$ is either real or imaginary. To put in a real form a formula in which $\Pi\mathrm{p}(u, a)$ is in fact imaginary, we write

$$\Pi\mathrm{p}(u, a) = i\Pi'\mathrm{p}(u, a);$$

if one of the two functions $\Pi\mathrm{p}(u, a)$, $\Pi'\mathrm{p}(u, a)$ is imaginary, the other is real. This notation is extremely convenient for our purpose here, but is obviously not susceptible of extension for general use.

11. The three functions $\mathrm{cs}^2 u$, $\mathrm{ds}^2 u$, $\mathrm{ns}^2 u$ are real on the perimeter $SCDNS$, and decrease steadily from $+\infty$ to $-\infty$ as u describes the contour; $\mathrm{cs}^2 u$ changes sign at C, $\mathrm{ds}^2 u$ at D, and $\mathrm{ns}^2 u$ at N. Hence $\mathrm{ps}\, a\ \mathrm{ps}'\, a$, which is $-\mathrm{cs}\, a\ \mathrm{ds}\, a\ \mathrm{ns}\, a$, is real if a is on SC or DN, imaginary if a is on CD or NS. It follows that $\Pi\mathrm{s}(u, a)$ is real if a is on SC or DN, and $\Pi'\mathrm{s}(u, a)$ is real if a is on CD or NS. We identify the side to which a belongs by reference to the value of one of the functions $\mathrm{pq}^2 a$; most simply $\mathrm{ds}^2 a$ decreases from $+\infty$ through c' to 0 along SCD and from 0 through $-c$ to $-\infty$ along DNS.

To locate a on a side of the fundamental rectangle by means of a real variable, we write $a = K_p + b$ or $a = K_p + ib'$, where K_p is one of the two corners available, and we have four pairs of formulae:

$$\mathrm{ds}^2 a > c'$$

(11.11) $a = b$ $\qquad \Pi\mathrm{s}(u, a) = \Pi\mathrm{s}(u, b) = \Lambda_s(u, b) + u\, \mathrm{zs}\, b$

(11.12) $a = K - b$ $\qquad \Pi\mathrm{s}(u, a) = -\Pi\mathrm{c}(u, b) = -\Lambda_c(u, b) - u\, \mathrm{zc}\, b,$

$$c' > \mathrm{ds}^2 a = 0$$

(11.13) $a = K + ib'$ $\qquad i\Pi'\mathrm{s}(u, a) = i\Pi'\mathrm{c}(u, ib') = \Lambda_c(u, ib') + u\, \mathrm{zc}\, ib'$

(11.14) $a = K + iK' - ib',$

$$i\Pi'\mathrm{s}(u, a) = -i\Pi'\mathrm{d}(u, ib') = -\Lambda_d(u, ib') - u\, \mathrm{zd}\, ib',$$

$$0 > \mathrm{ds}^2 a > -c$$

(11.15) $a = K + iK' - b$

$$\Pi\mathrm{s}(u, a) = -\Pi\mathrm{d}(u, b) = -\Lambda_d(u, b) - u\, \mathrm{zd}\, b$$

(11.16) $a = iK' + b$ $\qquad \Pi\mathrm{s}(u, a) = \Pi\mathrm{n}(u, b) = \Lambda_n(u, b) + u\, \mathrm{zn}\, b,$

$$-c > \mathrm{ds}^2 a$$

(11.17) $a = iK' - ib'$

$$i\Pi'\mathrm{s}(u, a) = -i\Pi'\mathrm{n}(u, ib') = -\Lambda_n(u, ib') - u\, \mathrm{zn}\, ib'$$

(11.18) $a = ib'$ $\qquad i\Pi'\mathrm{s}(u, a) = i\Pi'\mathrm{s}(u, ib') = \Lambda_s(u, ib') + u\, \mathrm{zs}\, ib'.$

For any one value of a there is a choice between two formulae, and we can cover the whole perimeter either using two theta functions with b, b' in the intervals $(0, K)$, $(0, K')$ or using the four theta functions with b, b' in the intervals $(0, \frac{1}{2}K)$, $(0, \frac{1}{2}K')$; in the first case we have a further choice, for we can use $\vartheta_s u$ on CSN and $\vartheta_d u$ on CDN or $\vartheta_c u$ on SCD and $\vartheta_n u$ on SND.

With a on SC or ND the choice between functions is more apparent than real. Writers from Legendre onwards ignore (11.12) and (11.15) without explaining why these alternatives can be ignored. For the final evaluation from (11.11) and (11.12) we have explicitly

$$\Lambda_s(u, b) = \tfrac{1}{2}\log\frac{\vartheta_s(b-u)}{\vartheta_s(b+u)}, \qquad \mathrm{zs}\, b = \frac{\vartheta_s' b}{\vartheta_s\, b},$$

$$\Lambda_c(u, b) = \tfrac{1}{2}\log\frac{\vartheta_c(b-u)}{\vartheta_c(b+u)}, \qquad \mathrm{zc}\, b = \frac{\vartheta_c' b}{\vartheta_c\, b}.$$

Since $\vartheta_s(K-u) = \vartheta_s K\, \vartheta_c u$, tables of $\vartheta_s u$ and $\vartheta_s' u$ have only to be provided with the complementary argument $K-u$ to become tables of $\vartheta_s K\, \vartheta_c u$ and $-\vartheta_s K\, \vartheta_c' u$, and we use the same entries and do the same arithmetic whether we compute $\Lambda_s(u, K-b)$ and $\mathrm{zs}(K-b)$ as

$$\tfrac{1}{2}\log\frac{\vartheta_s(K-b-u)}{\vartheta_s(K-b+u)}, \qquad \frac{\vartheta_s'(K-b)}{\vartheta_s\,(K-b)}$$

or compute $-\Lambda_c(u, b)$ and $-\mathrm{zs}\, b$ as

$$-\tfrac{1}{2}\log\frac{\vartheta_s K\vartheta_c(b+u)}{\vartheta_s K\vartheta_c(b-u)}, \qquad -\frac{\vartheta_s K\vartheta_c' b}{\vartheta_s K\vartheta_c\, b}.$$

The same considerations apply to (11.15) and (11.16): tables of $\vartheta_n u$ and $\vartheta_n' u$ provided with the complementary argument $K-u$ are tables of $\vartheta_n K\, \vartheta_d u$ and $-\vartheta_n K\, \vartheta_d' u$.

With a on SN or CD the process of evaluation is more elaborate and the distinction between the alternatives is not trivial. The theta function in $\Lambda_p(u, ib')$ has the complex arguments $ib' \pm u$ and must be dissected before $\Pi'\mathrm{p}(u, ib')$ can be computed. We take the four functions in turn. The theta functions are defined in terms of v, where $v/\frac{1}{2}\pi = u/K$, that is, where $v = \pi u/2K$ and we write also $\beta = \pi b'/2K$. It is to be noticed that $\vartheta_p' ib'$ means $(d\vartheta_p/du)_{u=ib'}$ that is, $(d\vartheta_p/dv)_{u=ib'} \,.\, dv/du$, and that therefore

$$u\,\vartheta_p' ib' = v(d\vartheta_p/dv)_{v=i\beta}.$$

The functions are defined in terms of v and q, where

$$q = e^{-\pi K'/K}, \tag{11.21}$$

but q is a constant of the Jacobian system and variation of q is not contemplated.

The functions $\vartheta_s u$, $\vartheta_c u$ are multiples of

$$\sin v - q^{1.2}\sin 3v + q^{2.3}\sin 5v - q^{3.4}\sin 7v + \ldots$$

$$\cos v + q^{1.2}\cos 3v + q^{2.3}\cos 5v + q^{3.4}\cos 7v + \ldots$$

and therefore $\vartheta_s(ib' + u)$ is a multiple of

$$(\cosh\beta\sin v - q^{1.2}\cosh 3\beta\sin 3v + q^{2.3}\cosh 5\beta\sin 5v - \ldots)$$
$$+\, i\,(\sinh\beta\cos v - q^{1.2}\sinh 3\beta\cos 3v + q^{2.3}\sinh 5\beta\cos 5v - \ldots)$$

and $\vartheta_c(ib' + u)$ is a multiple of

$$(\cosh\beta\cos v + q^{1.2}\cosh 3\beta\cos 3v + q^{2.3}\cosh 5\beta\cos 5v + \dots)$$
$$- i\,(\sinh\beta\sin v + q^{1.2}\sinh 3\beta\sin 3v + q^{2.3}\sinh 5\beta\sin 5v + \dots).$$

Hence

(11.22) $\Pi'\mathrm{s}(u, ib') =$

$$\arc\tan\frac{\cosh\beta\sin v - q^{1.2}\cosh 3\beta\sin 3v + q^{2.3}\cosh 5\beta\sin 5v\dots}{\sinh\beta\cos v - q^{1.2}\sinh 3\beta\cos 3v + q^{2.3}\sinh 5\beta\cos 5v\dots}$$
$$- u\,.\,\frac{\cosh\beta - 3q^{1.2}\cosh 3\beta + 5q^{2.3}\cosh 5\beta - \dots}{\sinh\beta - q^{1.2}\sinh 3\beta + q^{2.3}\sinh 5\beta - \dots},$$

and

(11.23) $\Pi'\mathrm{c}(u, ib\ ) =$

$$\arc\tan\frac{\sinh\beta\sin v + q^{1.2}\sinh 3\beta\sin 3v + q^{2.3}\sinh 5\beta\sin 5v + \dots}{\cosh\beta\cos v + q^{1.2}\cosh 3\beta\cos 3v + q^{2.3}\cosh 5\beta\cos 5v + \dots}$$
$$- u\,.\,\frac{\sinh\beta + 3q^{1.2}\sinh 3\beta + 5q^{2.3}\sinh 5\beta + \dots}{\cosh\beta + q^{1.2}\cosh 3\beta + q^{2.3}\cosh 5\beta + \dots}.$$

Similarly, since $\vartheta_n u$, $\vartheta_d u$ are multiples of

$$1 - 2q\cos 2v + 2q^4\cos 4v - 2q^9\cos 6v + 2q^{16}\cos 8v - \dots$$
$$1 + 2q\cos 2v + 2q^4\cos 4v + 2q^9\cos 6v + 2q^{16}\cos 8v + \dots$$

we have

(11.24) $\Pi'\mathrm{n}(u, ib') =$

$$\arc\tan\frac{2q\sinh 2\beta\sin 2v - 2q^4\sinh 4\beta\sin 4v + 2q^9\sinh 6\beta\sin 6v - \dots}{1 - 2q\cosh 2\beta\cosh 2v + 2q^4\cosh 4\beta\cos 4v - 2q^9\cosh 6\beta\cos 6v + \dots}$$
$$+ u\,.\,\frac{4q\sinh 2\beta - 8q^4\sinh 4\beta + 12q^9\sinh 6\beta - \dots}{1 - 2q\cosh 2\beta + 2q^4\cosh 4\beta - 2q^9\cosh 6\beta + \dots},$$

(11.25) $\Pi'\mathrm{d}(u, ib') =$

$$- \arc\tan\frac{2q\sinh 2\beta\sin 2\nu + 2q^4\sinh 4\beta\sin 4\nu + 2q^9\sinh 6\beta\sin 6\nu + \dots}{1 + 2q\cosh 2\beta\cos 2\nu + 2q^4\cosh 2\beta\cos 4\nu + 2q^9\cosh 6\beta\cos 6\nu + \dots}$$
$$- u\,.\,\frac{4q\sinh 2\beta + 8q^4\sinh 4\beta + 12q^9\sinh 6\beta + \dots}{1 + 2q\cosh 2\beta + 2q^4\cosh 4\beta + 2q^9\cosh 6\beta + \dots}.$$

If b' and u are real, the functions $\Pi'\mathrm{p}(u, ib')$ have real values and (11.22)—(11.25) are formulae from which these values can be calculated. The hyperbolic functions do not retard appreciably the convergence of the several series; if b' is in the range $(0, K')$, both $\sinh n\beta$ and $\cosh n\beta$ are smaller than q^{-n}, and if b' is in $(0, \frac{1}{2}K')$, then $\sinh 2n\beta$ and $\cosh 2n\beta$ are smaller than q^{-n}. The restriction on the path of u implies that the inverse tangents are all in the interval $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$.

The dissection of the theta functions for the evaluation of elliptic integrals is classical; the improvement on current practice lies in avoiding a mixture of functions in any one formula.

12. Light is thrown on the alternatives in (11.11)—(11.18) by the relation (6.2) between $\Pi\mathrm{s}(u, a)$ and $\Pi\mathrm{p}(u, a)$:

$$(12.1)\qquad \Pi\mathrm{p}(u, a) = \Pi\mathrm{s}(u, a) + \tfrac{1}{2}\log\frac{\mathrm{ps}(a-u)}{\mathrm{ps}(a+u)} + u\,.\,\frac{\mathrm{ps}'a}{\mathrm{ps}\,a}\,.$$

Denote the midpoints of SC, CD, DN, NS by E, F, G, H, and let $b \equiv b_s$ be a point in SE and $ib' \equiv b'_s$ be a point in SH.

In the half-sides EC, DG, GN there are points b_c, b_d, b_n at distance b from the corners C, D, N, and we have

$$\begin{aligned} b_c &= K_c - b_s, &\quad \Pi\mathrm{s}(u, b_c) &= -\,\Pi\mathrm{c}(u, b_s),\\ b_d &= -K_d - b_s, &\quad \Pi\mathrm{s}(u, b_d) &= -\,\Pi\mathrm{d}(u, b_s),\\ b_n &= K_n + b_s, &\quad \Pi\mathrm{s}(u, b_n) &= \Pi\mathrm{n}(u, b_s). \end{aligned}$$

If b_s traverses SE, the four points b_s, b_c, b_n, b_d together traverse the two sides SC, ND, and the evaluation of $\Pi\mathrm{s}(u, a)$ is extended from SE to the two sides by means of the elliptic functions ps u:

$$(12.21)\qquad \Pi\mathrm{s}(u, b_s) = \Pi\mathrm{s}(u, b)$$

$$(12.22)\qquad \Pi\mathrm{s}(u, b_c) = -\,\Pi\mathrm{s}(u, b) - \tfrac{1}{2}\log\frac{\mathrm{cs}(b-u)}{\mathrm{cs}(b+u)} - u\,.\,\frac{\mathrm{cs}'b}{\mathrm{cs}\,b}$$

$$(12.23)\qquad \Pi\mathrm{s}(u, b_n) = \Pi\mathrm{s}(u, b) + \tfrac{1}{2}\log\frac{\mathrm{ns}(b-u)}{\mathrm{ns}(b+u)} + u\,.\,\frac{\mathrm{ns}'b}{\mathrm{ns}\,b}$$

$$(12.24)\qquad \Pi\mathrm{s}(u, b_d) = -\,\Pi\mathrm{s}(u, b) - \tfrac{1}{2}\log\frac{\mathrm{ds}(b-u)}{\mathrm{ds}(b+u)} - u\,.\,\frac{\mathrm{ds}'b}{\mathrm{ds}\,b}\,.$$

Since the operation of evaluating the difference

$$\tfrac{1}{2}\log\frac{\mathrm{ps}(b-u)}{\mathrm{ps}(b+u)} + u\,.\,\frac{\mathrm{ps}'b}{\mathrm{ps}\,b}$$

from tables of ps u and ps$'$ u is precisely the same as the operation of evaluating $\Pi\mathrm{p}(u, a)$ in the form

$$\tfrac{1}{2}\log\frac{\vartheta_p(a-u)}{\vartheta_p(a+u)} + u\frac{\vartheta_p'a}{\vartheta_p a}$$

from tables of $\vartheta_p u$ and $\vartheta_p'u$, no practical advantage is to be expected from these formulae.

It is different when we deal with the half-sides HN, CF, FD. On them we have points b'_n, b'_c, b'_d such that

$$\begin{aligned} b'_n &= K_n - b'_s, &\quad \Pi\mathrm{s}(u, b'_n) &= -\,\Pi\mathrm{n}(u, b'_s),\\ b'_c &= K_c + b'_s, &\quad \Pi\mathrm{s}(u, b'_c) &= \Pi\mathrm{c}(u, b'_s),\\ b'_d &= -K_d - b'_s, &\quad \Pi\mathrm{s}(u, b'_d) &= -\,\Pi\mathrm{d}(u, b'_s). \end{aligned}$$

Since b'_s is imaginary, we take (6.2) in the form

$$\text{(12.3)} \qquad i\Pi'\mathrm{p}(u, a) = i\Pi'\mathrm{s}(u, a) + \tfrac{1}{2}\log\frac{\mathrm{bps}\, a + \mathrm{bps}\, u}{\mathrm{bps}\, a - \mathrm{bps}\, u} - u\,\mathrm{bps}\, a.$$

Using Jacobi's imaginary transformation we have

$$\mathrm{bcs}(ib'|c) = i\,\mathrm{bns}(b'|c'),\ \mathrm{bds}(ib'|c) = i\,\mathrm{bds}(b'|c'),\ \mathrm{bns}(ib'|c) = i\,\mathrm{bcs}(b'|c')$$

and therefore

$$\text{(12.41)} \qquad \Pi'\mathrm{s}(u, b'_s) = \Pi'\mathrm{s}(u, ib'),$$

$$\text{(12.42)} \qquad \Pi'\mathrm{s}(u, b'_n) = -\Pi'\mathrm{s}(u, ib') + \arctan\frac{\mathrm{bcs}(b'|c')}{\mathrm{bns}(u|c)} - u\,\mathrm{bcs}(b'|c'),$$

$$\text{(12.43)} \qquad \Pi'\mathrm{s}(u, b'_c) = \Pi'\mathrm{s}(u, ib') - \arctan\frac{\mathrm{bns}(b'|c')}{\mathrm{bcs}(u|c')} + u\,\mathrm{bns}(b'|c'),$$

$$\text{(12.44)} \qquad \Pi'\mathrm{d}(u, b'_d) = \Pi'\mathrm{s}(u, ib') + \arctan\frac{\mathrm{bds}(b'|c')}{\mathrm{bds}(u|c)} - u\,\mathrm{bds}(b'|c').$$

It is far quicker to evaluate a difference

$$\arctan\frac{\mathrm{bqs}(b'|c')}{\mathrm{bps}(u|c)} - u\,\mathrm{bqs}(b'|c')$$

than to find an isolated value of a function $\Pi'\mathrm{p}(u, ib')$ by means of a dissected q-series, and (12.41)—(12.44), unlike (12.21)—(12.24), can be recommended to computers.

13. To conclude, we have to consider the integral

$$L = \int_{u_1}^{u_2} \frac{du}{1 - \mu\,\mathrm{pq}^2 u}$$

between arbitrary real limits. If the integral can be expressed as the difference between integrals from 0, the evaluation in one of the forms

$$\frac{\mathrm{sp}\, a}{\mathrm{sp}'a}\Pi_{12},\ (u_2 - u_1) + \frac{\mathrm{ps}\, a}{\mathrm{ps}'a}\Pi_{12},\ \frac{u_2 - u_1}{1 - \mu} + \frac{\mathrm{qp}\, a}{\mathrm{qp}'a}\Pi_{12},$$

where $\Pi_{12} = \Pi\mathrm{s}(u_2, a) - \Pi\mathrm{s}(u_1, a)$ introduces no fresh problems. But since the integral has a logarithmic singularity at any point where $\mathrm{pq}^2u = \mathrm{pq}^2a$, there is a tacit assumption throughout that there is no such point on the u-path.

If a is not real, this assumption does not come into operation. But if a is real, $\Pi\mathrm{s}(u, a)$ is defined as a real integral only for values of u in $(-a, a)$ and L is expressible by means of Π_{12} only if u and u_2 are in this interval, whereas the condition implicit in the existence of the integral does not restrict u and u_2 separately. The problem is the same as in the integration of $1/x$. If neither $\vartheta_s(a - u)$ nor $\vartheta_s(a + u)$ is zero for any value of u in (u_1, u_2), the two quotients $\vartheta_s(a - u_2)/\vartheta_s(a - u_1)$, and $\vartheta_s(a + u_1)/\vartheta_s(a + u_2)$ are positive and Π_{12}, defined as

$$\int_{u_1}^{u_2} J_s(u, a)\, du,$$

can be compu ted as

$$\tfrac{1}{2}\log\frac{\vartheta_s(a-u_2)\vartheta_s(a+u_1)}{\vartheta_s(a-u_1)\vartheta_s(a+u_2)} + (u_2-u_1)\frac{\vartheta_s'a}{\vartheta_s a}.$$

If there are points $b_1, b_2, \ldots, b_m$ in (u_1, u_2) such that $\mathrm{qp}^2 b_r = \mathrm{qp}^2 a$ the substitution of

$$\tfrac{1}{2}\log\left|\frac{\vartheta_s(a-u_2)\vartheta_s(a+u_1)}{\vartheta_s(a-u_1)\vartheta_s(a+u_2)}\right| + (u_2-u_1)\frac{\vartheta_s'a}{\vartheta_s a}$$

for Π_{12}, in the formal evaluation gives the limit of the sum

$$\int_{u_1}^{b_1-\epsilon_1} + \int_{b_1+\epsilon_1}^{b_2-\epsilon_2} + \int_{b_{m-1}+\epsilon_{m-1}}^{b_m-\epsilon_m} + \int_{b_m+\epsilon_m}^{u_2} \frac{du}{1-\mu\,\mathrm{pq}^2 u}$$

when $\epsilon_1, \epsilon_2, \ldots, \epsilon_m$ tend independently to zero.

References

1. C. A. A. Briot et J. C. Bouquet, *Théorie des Fonctions elliptiques*, (Paris, 1875).
2. A. Enneper, *Elliptische Funktionen, Theorie und Geschichte* (Halle, 1876).
3. C. Hermite, *Note sur la théorie des fonctions elliptiques* in Serret, *Cours de calcul différentiel et intégral* (4th ed, Paris, 1894) 737–904.
4. C. G. J. Jacobi, *Fundamenta nova theoriae functionum ellipticarum* (Königsburg, 1829).
5. A. M. Legendre, *Exercices de calcul intégral* (Paris, 1811).
6. ——— *Théorie des fonctions elliptiques*, t.1 (Paris, 1825).
7. H. Lenz, *Uber die elliptischen Funktionen von Jacobi*, Math. Zeit., *67* (1957), 153–175.
8. E. H. Neville, *Jacobian Elliptic Functions* (2nd ed., Cambridge, 1951).

Sonning-on-Thames,
England

MIXED PROBLEMS FOR HYPERBOLIC EQUATIONS OF GENERAL ORDER

G. F. D. DUFF

The object of this paper is the extension to linear partial differential equations of order m in N independent variables, of the existence theorems for mixed initial and boundary value problems which have been established for systems of first order equations in **(3)**. In such mixed problems an initial surface S and a boundary surface T are the carriers of the two types of data, and the number of datum functions to be assigned on T depends on the configuration of the characteristic surfaces relative to S and T.

For the first part of the paper (§§ 1-5) the coefficients in the differential equation, the initial and boundary surfaces, and the data prescribed are all taken to be real analytic in the variables $x^1 \ldots x^N$. In this "analytic" case an existence theorem is established for boundary conditions of considerable generality. We assume that the differential equation is regularly hyperbolic with respect to S and T, a notion which is stated precisely in § 1, and is weaker than the usual regular hyperbolic condition. Then the single equation of higher order is reduced to a system of equations of first order, of the type treated in **(3)**, and the existence theorem there established is taken over to obtain the result, which is stated as Theorem 1 in § 5 below. For this purpose we require a certain algebraic lemma relating to the characteristic roots.

The non-analytic problem for regularly hyperbolic equations is treated in §§ 6-10, by adaptation of the energy integral method. A general sufficient condition for the existence of a solution is given in § 6. As it appears that this condition is not always fulfilled, it is necessary to discuss particular cases. In § 8 and § 9 are treated two such special problems, each of which is a generalization of the known results for second order equations. The first of these concerns the problem wherein the number of boundary conditions is one less than the number of initial conditions. The second requires an assumption of symmetry relative to the boundary surface, and the number of boundary conditions is half the number of initial conditions.

1. The differential equation. We consider an analytic linear partial differential equation of order m in the N independent variables x^i:

$$(1.1) \qquad Lu = \sum_{h=0}^{m} a_{(h)i_1\ldots i_h} \frac{\partial^h u}{\partial x^{i_1} \ldots \partial x^{i_h}} = 0.$$

The dependent variable is $u = u(x^i)$. The coefficients

$$a_{(h)i_1\ldots i_h}$$

Received May 3, 1958.

of order h are assumed in this first part to be convergent real power series of the real variables x^i.

Let $S : \phi(x^i) = 0$ be an "initial" surface not characteristic for the linear operator L; and let $T : \psi(x^i) = 0$ be a "boundary" surface likewise not characteristic. We assume that both S and T are analytic and that they have an $(N - 2)$-dimensional intersection C.

The characteristic surfaces $G : \chi(x^i) = 0$ of the operator L satisfy the equation

$$C[\chi] = a_{(m)i_1 \dots i_m} \frac{\partial \chi}{\partial x^{i_1}} \cdots \frac{\partial \chi}{\partial x^{i_m}} = 0, \tag{1.2}$$

and in general there will be m (or fewer) characteristic surfaces which pass through the edge C. As seen below we assume that there are actually m. We suppose that at least $k_0(k_0 < m)$ of these lie in a fixed "quadrant" R defined by S and T: and we select k_0 of these surfaces G_i, $(i = 1, \dots, k_0)$. These shall be referred to as "select" characteristic surfaces, and all others as "non-select."

The mixed problem to be studied below is now formulated as follows. Define $t = \phi(x^i)$, $x = \psi(x^i)$, and assign on S Cauchy data for u with respect to the operator L: that is, values of u and its derivatives with respect to t up to order $m - 1$ inclusive. Assign on T any k_0 of the m quantities:

$$u, \frac{\partial u}{\partial x}, \dots, \frac{\partial^{m-1} u}{\partial x^{m-1}},$$

subject to compatibility conditions of order $m - 1$ on the edge C. We seek a piecewise analytic solution in R of $Lu = 0$ which takes the given values on S and on T, and is analytic except across the select characteristic surfaces, where it is of class C^{m-1}.

In order to treat this problem we shall need to assume that the operator L is regularly hyperbolic with respect to S and T, in the following sense: *there shall be m distinct characteristic surfaces passing through C.* Another form of this condition is available if we consider the coefficients

$$a_{(m)i_1 \dots i_m}$$

of highest order derivatives in Lu, in which the indices i_p take values N and $N - 1$ corresponding to t and x, respectively. To find this condition we note that by the theory of first order partial differential equations, the characteristic surfaces through C are composed of characteristic strips of the reduced characteristic equation

$$C_1[p_\rho, -1] = a_{(m)i_1 \dots i_m} p_{i_1} \dots p_{i_m} = 0, \tag{1.3}$$

where $p_N \equiv p_t \equiv -1$ has been substituted so that χ appears in the form $\chi = \chi_0(x_\rho, x) - t$. On the initial "curve" C the initial values for the strip elements are found from (1.3) and the conditions

$$dt = \sum_{\rho=1}^{N-1} p_\rho dx_\rho.$$

Since for $\rho = 1, 2, \ldots, N-2$, the dx_ρ are independent, we have $p_\rho = 0$, $(\rho = 1, 2, \ldots, N-2)$, and only p_x is different from zero. Thus

$$C_1[0, 0, \ldots, 0, p_x, -1] = 0 \tag{1.4}$$

determines the m values of p_x, which we shall suppose are real and distinct on C and in a neighbourhood of C. Setting

$$a_k = a_{(m)N-1, N-1, \ldots, N-1, N, N, \ldots, N},$$

where the index $N-1$ appears k times, we see that (1.4) becomes

$$a_m p_x^m - a_{m-1} p_x^{m-1} + \ldots + (-1)^m a_0 = 0. \tag{1.5}$$

That is, the roots p_x of (1.5) must be distinct and real.

We note that if $Lu = 0$ is regularly hyperbolic in the sense of Leray **(8)** and if S is spacelike, then Lu is regularly hyperbolic with respect to S and to every non-characteristic surface T. If in the regularly hyperbolic case we imagine the edge C to rotate about a fixed point in S, the characteristic surfaces issuing from C remain separated: no two touch. For our purpose it is enough if these surfaces are distinct for the one position of the edge C. Thus our condition is weaker than the customary regular hyperbolic condition. In fact it becomes equivalent for the case of two independent variables, when the edge C reduces to a point.

Analogously, the normal surface of a regularly hyperbolic operator consists essentially of a nest of concentric ovals, such that any line through the origin meets the surface in a maximal number of real points. In our case it is sufficient if a particular line through the origin, namely the normal to C in S, meets the surface in a maximal number of real points. This could be realized, for instance, by a surface with multiple points, or by a surface consisting of ovals external to one another.

If we alter the negative signs in (1.5) and consider the equation

$$a_m \gamma^m + a_{m-1}\gamma^{m-1} + \ldots + a_0 = 0, \tag{1.6}$$

we see that the roots $\gamma_1, \ldots, \gamma_m$ of (1.6) are also distinct. Let us define the first order operator

$$D_s u = \frac{\partial u}{\partial x} + \gamma_s \frac{\partial u}{\partial t}, \tag{1.7}$$

which indicates differentiation along the section of a characteristic surface by a plane $x_\rho = \text{const}$, $(\rho = 1, \ldots, N-2)$.

The operators D_i shall also be termed select or non-select according as the characteristic surface G_i and the characteristic root γ_i are select or not. We note the identity

$$a_m \prod_s D_s u = \sum_{k=0}^{m} a_k \frac{\partial^m u}{\partial x^k \partial t^{m-k}} + \ldots \tag{1.8}$$

where the terms omitted are derivatives of order less than m. In consequence, we can write the given differential equation (1.1) in the form

$$\prod_{a} D_a u = L_1(u) \tag{1.9}$$

where $L_1(u)$ is a linear operator of order m in which no term has more than $m - 1$ differentiations with respect to x and t combined.

2. Reduction to a system of first order equations. By introducing as new dependent variables v_i suitable combinations of the derivatives of u, we shall perform a formal reduction of (1.1) to a system

$$D_a v_i = \sum c_{ij} \frac{\partial v_i}{\partial x^j} + \sum e_{ij} v_j, \tag{2.1}$$

where each D_a operator will occur several times, in general, and where only derivatives with respect to x_ρ $(\rho = 1, \ldots, n-2)$ appear as $\partial/\partial x_\rho$ on the right. This system is of the type studied in (3), as the elementary divisors of A^N relative to A^{N-1} are simple. The dependent variables in (2.1) are also divided into select and non-select classes, a variable v being select if the operator D_a which operates on it in (2.1) is select, and vice versa. The assigned boundary conditions will be transformed into

$$v_i = \sum a_{ij} v_j + f_i \tag{2.2}$$

where on the left shall appear only the select v_i. Thus the existence theorem of **(3)** is applicable and will lead to a solution of the original problem, when that problem is analytic.

The formal reduction and labelling of new variables will follow the pattern of the Cauchy-Kowalewski reduction to normal form, except for those derivatives with respect to x. To handle these we have introduced the D_a operators and will employ them in the fashion of **(1)**. A result of this distinction is the following subdivision of the new variables into groups. We define formally

$$v_{(i)} = v_{ab\ldots h\, i_1\ldots i_q} = D_a D_b \ldots D_h \frac{\partial^q u}{\partial x^{i_1} \ldots \partial x^{i_q}} \tag{2.3}$$

and construct the first order system so as to satisfy these definitions identically.

In the first group we take $q = 0$; and for higher values of q up to $m - 1$ inclusive there is a group of equations corresponding to each distinct array $i_1, \ldots, i_q$, where order is immaterial.

A fixed ordering $a, b, c, \ldots, k, \ldots$ of the operators D_a is used in each of these groups. However, as the exact selection of these indices depends on the boundary conditions, we shall for the present reserve the choice of select and non-select values.

Certain "commutator" expressions appear and we now define them. The symbol

$$C_{a\, i_q b\ldots hk\; i_1\ldots i_{q-1}}[u] \tag{2.4}$$

shall denote the reduced form of the expression

$$D_a \frac{\partial}{\partial x^{i_q}} D_b \ldots D_h D_k \frac{\partial^{q-1}u}{\partial x^{i_1} \ldots \partial x^{i_{q-1}}} - \frac{\partial}{\partial x^{i_q}} D_k D_h \ldots D_a \frac{\partial^{q-1}u}{\partial x^{i_1} \ldots \partial x^{i_{q-1}}}; \tag{2.5}$$

this reduced form contains derivatives of u of lower orders $a, b, \ldots, h, k, i_1 \ldots i_{q-1}$, with coefficients functions of the γ_a.

By writing

$$C_{a i_q b \ldots hk i_1 \ldots i_{q-1}}[v]$$

we shall indicate that the derivatives of u have been formally replaced by the corresponding variables

$$v_{ab \ldots h i_1 \ldots i_q}$$

as defined in (2.3). This is possible, since we will show that all derivatives of u can be expressed as linear combinations of the variables $v_{(i)}$ in (2.3). In fact we shall prove by induction that all $k + 1$ derivatives

$$\frac{\partial^k u}{\partial x^h \partial t^{k-h}} \qquad h = 0, \ldots, k,$$

can be expressed as linear combinations of the $k + 1$ variables

$$v_{hg \ldots a}, \quad v_{kg \ldots a}, \quad \ldots, \quad v_{kh \ldots b},$$

in each of which one of the $k + 1$ operations is omitted, and of variables with a lesser number of subscripts. To show this, we note that by (1.7) and (2.3)

$$v_{kh \ldots b} = \sum_{h=0}^{k} S_h^a(\gamma) \frac{\partial^k n}{\partial x^h \partial t^{k-h}} + F_{k-1}[u] \tag{2.6}$$

where $S_h^a(\gamma)$ is the symmetric function of degree h of the k quantities γ_k, $\gamma_h, \ldots, \gamma_b$ and with γ_a omitted. Forming similar equations with $b, c, \ldots, h, k$ omitted in turn, we see that the system can be solved for the kth order derivatives of u provided that the determinant $|S_h^a(\gamma)|$ is not zero. This is proved separately in Lemma 1 below, and thus our assertion is verified.

In (2.6) $F_{k-1}[u]$ denotes an operator in $\partial/\partial x$, $\partial/\partial t$ of order $k - 1$, which by the induction hypothesis may be considered to be already expressed as a combination of the v's.

The groups shall be written with definitions and differential equations in parallel columns. For $q = 0$ we have the "triangular" array of equations shown in Table I.

In Table I $a, b, c, \ldots, k, \ldots, n$ denote the distinct numbers from 1 to n in an as yet undefined order. The operator $L_1[v]$ is defined by replacing derivatives of u in $L_1(u)$ (cf. (1.9)) with the appropriate first derivatives or values of the v's. This array or group of equations contains n subgroups, the kth group containing k equations each with a different operator D_a.

For each array $(i)_q = (i_1, \ldots, i_q)$ we have a similar group of equations, in which appear on the right side certain first derivatives with respect to x_q^i. We let $(i)_{q-1} = (i_1 \ldots i_{q-1})$ and construct Table II, which also contains a triangular array, with subgroups of increasing size.

TABLE I

DEFINITIONS	DIFFERENTIAL EQUATIONS
$v = u$	$D_a v = v_a$
$v_b = D_b u$	$D_a v_b = v_{ba} + C_{ab}[v]$
$v_a = D_a u$	$D_b v_a = v_{ba}$
$v_{cb} = D_c D_b u$	$D_a v_{cb} = v_{cba} + C_{acb}[v]$
$v_{ca} = D_c D_a u$	$D_b v_{ca} = v_{cba} + C_{bca}[v]$
$v_{ba} = D_b D_a u$	$D_c v_{ba} = v_{cba}$
	$\cdots$
$v_{kh\ldots b} = D_k D_h \ldots D_b u$	$D_a v_{kh\ldots b} = v_{khg\ldots a} + C_{akh\ldots b}[v]$
	$\cdot$
$v_{kg\ldots a} = D_k D_g \ldots D_a u$	$D_h v_{kg\ldots a} = v_{khg\ldots a} + C_{hkg\ldots a}[v]$
$v_{hg\ldots a} = D_h D_g \ldots D_a u$	$D_k v_{hg\ldots a} = v_{khg\ldots a}$
	$\cdots$
$v_{nm\ldots b} = D_n D_m \ldots D_b u,$	$D_a v_{nm\ldots b} = L_1[v] + C_{anm\ldots b}[v]$
	$\cdot$
$v_{n1\ldots a} = D_n D_1 \ldots D_a u,$	$D_m v_{n1\ldots a} = L_1[v] + C_{nm1\ldots a}[v]$
$v_{m1\ldots a} = D_m D_1 \ldots D_a u,$	$D_n v_{m1\ldots a} = L_1[v].$

The last of the "subgroups" of Table II contains $m - q$ equations, and the order of $a, b, c \ldots h, k$ is the same in all groups, that is, for all $(i)_q$. It is seen that the number of new variables defined in these groups is equal to the total number of partial derivatives of u with respect to t, x and the x_p, up to and including order $m - 1$. We emphasize that not all groups $(i_1 \ldots i_q)$ are here represented, but only one for each set of integers $(i_1 \ldots i_q)$ without regard to order. Thus we may for simplicity assume that $i_1 < i_2 < i_3 < \ldots < i_q$.

3. Reduction of the boundary conditions. Let the derivatives of u with respect to x up to order $n - 1$ be paired in order with the numbers $a, b, c, \ldots, k$, which label the D_a operators:

$$(3.1)\qquad \begin{array}{ccccc} u & u_x & u_{xx} & , \ldots & u_x^{(n-1)} \\ a & b & c & , \ldots & l. \end{array}$$

We establish an ordered correspondence between the operators of the sequence $D_a, D_b, \ldots, D_k$ in Table I and the derivatives $u, u_x, \ldots, u_x^{n-1}$. The labels $a, b, \ldots, k$ shall be chosen so that to an assigned derivative $u_x^{(j)}$ there corresponds a select operator D_h, and vice versa. This is possible, in general in a number of ways, since the number of select operators has been taken as equal to the number of boundary conditions. It is now assumed that this arrangement has been adopted in advance in Tables I and II. We repeat that the select v's are those which are operated on in (2.1) by a D_a operator which is select according to this scheme.

TABLE II

DEFINITIONS	DIFFERENTIAL EQUATIONS
$v_{(t)} = \dfrac{\partial^q u}{\partial x^{t_1} \ldots \partial x^{t_q}} = u_{(t)}$,	$D_a v_{(t)} = \dfrac{\partial}{\partial x_{t_q}} v_{a(t)_{q-1}} + C_{a t_1 \ldots t_q}[v]$
$\ldots$	
$v_{b(t)} = D_b u_{(t)}$,	$D_a v_{b(t)} = \dfrac{\partial}{\partial x^{t_q}} v_{ba(t)_{q-1}} + C_{a b t_1 \ldots t_q}[v]$
$v_{a(t)} = D_a u_{(t)}$,	$D_b v_{a(t)} = \dfrac{\partial}{\partial x^{t_q}} v_{ba(t)_{q-1}} + C_{b t_q a t_1 \ldots t_{q-1}}[v]$
$\ldots$	
$v_{cb(t)} = D_c D_b u_{(t)}$,	$D_a v_{cb(t)} = \dfrac{\partial}{\partial x^{t_q}} v_{cba(t)_{q-1}} + C_{a t_q c b t_1 \ldots t_{q-1}}[v]$
$v_{ca(t)} = D_c D_a u_{(t)}$,	$D_b v_{ca(t)} = \dfrac{\partial}{\partial x^{t_q}} v_{cba(t)_{q-1}} + C_{b t_q c a t_1 \ldots t_{q-1}}[v]$
$v_{ba(t)} = D_b D_a u_{(t)}$,	$D_c v_{ba(t)} = \dfrac{\partial}{\partial x^{t_q}} v_{cba(t)_{q-1}} + C_{c t_q b a t_1 \ldots t_{q-1}}[v]$
$\ldots$	
$\ldots$	
$v_{kh \ldots b(t)} = D_k D_h \ldots D_b u_{(t)}$,	$D_a v_{kh \ldots b(t)} = \dfrac{\partial}{\partial x^{t_q}} v_{khg \ldots a(t)_{q-1}} + C_{a t_q kh \ldots b t_1 \ldots t_{q-1}}[v]$
$\cdot$	
$v_{kg \ldots a(t)} = D_k D_g \ldots D_a u_{(t)}$,	$D_h v_{kg \ldots a(t)} = \dfrac{\partial}{\partial x^{t_q}} v_{khg \ldots a(t)_{q-1}} + C_{h t_q kg \ldots a t_1 \ldots t_{q-1}}[v]$
$v_{hg \ldots a(t)} = D_h D_g \ldots D_a u_{(t)}$,	$D_k v_{hg \ldots a(t)} = \dfrac{\partial}{\partial x^{t_q}} v_{khg \ldots a(t)_{q-1}} + C_{k t_q hg \ldots a t_1 \ldots t_{q-1}}[v]$

These boundary conditions must be reduced to the form (2.2) where the select v's appear only on the left. Let us consider first the group $q = 0$: for the other groups the calculations are similar. If u is given, D_a and v shall be select; and then $v = u$ is a boundary condition of type (2.2). If u is not given then no boundary condition for v will be needed, as D_a is then non-select.

If u_x is given then D_b is select. There are two cases, according as D_a is select or not. If D_a is select, then $v_a = u_x + \gamma_a u_t$ and $v_b = u_x + \gamma_b u_t$ are both assigned and these conditions are of the form (2.2). If D_a is not select, we eliminate u_t between these two relations and find

$$v_b = \frac{\gamma_b}{\gamma_a} v_a + \left(1 - \frac{\gamma_b}{\gamma_a}\right) u_x, \tag{3.2}$$

which again has the form (2.2). However, if u_x is not given, there are two other cases. If D_a is select, and D_b is not ,we can solve (3.2) for v_a, which is then the single necessary boundary condition of form (2.2). If neither D_a nor D_b is select, no boundary condition is needed.

We may now proceed by induction from one subgroup to the next. If the boundary conditions for the kth subgroup have been put in the form (2.2) then all select variables of that group can be expressed in terms of non-select variables. Thus in treating the next sub-group we can allow all variables of the preceding groups to appear on the right side of the boundary conditions, as the select ones can later be eliminated by means of the preceding boundary conditions. If in (2.6) we replace the quantity $F_{k-1}[u]$ by its formal equivalent in terms of the v's, we see that only those v's of the preceding groups will appear. Consequently this term of (2.6) may be considered as non-essential in the remainder of the calculation.

Assuming then that the result holds for the $(k-1)$st subgroup, let us prove it for the kth subgroup. We divide the k equations (2.6) into select and non-select categories according as the v variable on the left is select or not. All derivatives of $u, u_x, \ldots, u_x^{(s)}$ with respect to t are known, or select, on T, according as $u, u_x, \ldots, u_x^{(s)}$ is select or not. Thus if h of the k quantities $u, u_x, \ldots, u_x^{(k)}$ are select, then h of the derivatives written explicitly on the right side of (2.6) are select. Let us pick out the $k-h$ non-select equations of (2.6) and solve them for the $k-h$ non-select derivatives of u in terms of the select derivatives of u and the (non-select) variables v of these equations. The possibility of this depends on the non-vanishing of a determinant of which the elements are symmetric functions of the γ's, with one of the γ's omitted in each row. Supposing, as will be shown in Lemma 1 of § 4, that all such determinants are different from zero, we can carry out this inversion of the non-select equations (2.6), and then replace the $k-h$ non-select derivatives of u in the h select equations of (2.1) by the expressions so found for them. These h select equations will then take the form (2.2) since all earlier groups of select variables can be eliminated from the right sides.

Now consider these groups with $q > 0$. As all differentiations with respect to x_ρ $(\rho = 1, \ldots, N-2)$ are tangential to T, the derivatives

$$\partial^{q+h} u / \partial x^{i_1} \ldots \partial x^{i_q} \partial x^h$$

are select according as $\partial^h u / \partial x^h$ is select, or not. Thus equations similar to (2.6) can be written down for any such group, and the coefficients of the terms shown explicitly will be exactly the same, and the structure of the operator $F_{k-1}[u]$ will be unaltered except that instead of u as argument we will have

$$\partial^q u / \partial x^{i_1} \ldots \partial x^{i_q}.$$

It follows that the boundary conditions for all groups with $0 < q < m-1$ can be put in the same form (2.2).

Remark. Suppose that the k linear boundary conditions are linear and independent relations among the n quantities $u, u_x, u_{xx}, \ldots, u_x^{(n-1)}$, on T.

Such a system of boundary conditions can be reduced to a triangular standard form

$$(3.3) \qquad u_x^{(h_i)} + \sum_{r_i=0}^{h_i-1} c_{r_i} u_x^{(r_i)} = g_i, \qquad (i = 1, \ldots, k_0),$$

where the orders h_i of the leading derivatives form an increasing sequence:

$$h_1 < h_2 < \ldots < h_{k_0}.$$

In addition, the indices r_i are all less than h_i, while the coefficients c_{r_i} are analytic functions on T.

To show that this system of boundary conditions can be expressed in the form (2.2), we need modify the previous working only slightly. In the array (3.1) we choose as select operators $D_i, \ldots, D_n$ those corresponding to the

$$u_x^{(h_i)}.$$

We then commence with the quantity u: if and only if it is given in the first array $q = 0$ of Table I, a boundary condition is required. Now, considering u_x, we see that if one of the $h_i = 1$, there is a condition

$$u_x + c_{11}u = g_1,$$

and if u is given it may be replaced by its given values while, if it is not, the corresponding variable v of Table I is non-select and so may appear on the right of the boundary conditions (2.2). Thus the form (2.2) is attained in either case, as in the preceding calculations.

Proceeding by induction on h_i, we see that in the typical condition (3.3) the terms $u_x^{(r_i)}$ are either non-select, in which case they are allowed on the right side of (2.2), or else they can be expressed, by means of the boundary conditions already standardized, in terms of non-select variables and given data. As remarked earlier any variable of a previous group can be allowed on the right side of a boundary condition in the course of such a calculation. This completes the demonstration that (3.3) can be reduced to the standard form of the boundary conditions for the system of first order equations.

4. A lemma on symmetric functions. To justify the reduction of the differential equation as well as the boundary conditions, we establish a lemma which is required in its most general form in the preceding discussion of the boundary conditions.

LEMMA 1. *Let k distinct numbers $\gamma_a \neq 0$ be given, and let $s_r^a(\gamma) = s_r^a$ denote the elementary symmetric function of degree, r of all $k - 1$ γ's with γ_a omitted. Then every subdeterminant formed by deleting an equal number $h(0 < h < k - 1)$ of rows and columns from the $k \times k$ determinant*

$$(4.1) \qquad |s_r^a(\gamma)|,$$

is different from zero.

The numbers γ_a corresponding to the deleted rows are the select γ_a; for convenience we shall denote them now by c_a while retaining γ_a for the $l = k-h$ non-select numbers. The deleted columns refer to the assigned derivatives among $u, u_x, u_x, \ldots, u_x^{(m-1)}$.

Let $s_r(c)$ denote the elementary symmetric function of degree r of the c_a: if we now delete the h select rows we observe from (4.1), that all h of the select c_a are present in each of the other rows. Let $\sigma_r^a(\gamma) = \sigma_r^a$ denote the elementary symmetric function of degree r of the non-select γ's, with γ_a omitted. The following property of s_r^a is evident: s_r^a is the convolution of the $s_m(c)$ and $\sigma_{r-m}^a(\gamma)$ of all lower orders:

$$s_r^a(\gamma) = \sum_{m=0}^{r} s_m(c)\sigma_{r-m}^a(\gamma). \tag{4.2}$$

Now let σ_r denote the column vector with $l = k - h$ components σ_r^a. The $k \times (l - h)$ matrix (the select rows have been deleted), may now be written

$$\left(\sum_{i=0}^{k-1} s_i\sigma_{k-i-1} \right).$$

We note that $\sigma_r = 0$ for $r > l$, so that we may write this array in the form

$$(s_0\sigma_0, s_0\sigma_1 + s_1\sigma_0, s_0\sigma_2 + s_1\sigma_1 + s_2\sigma_0, \ldots, s_{l-1}\sigma_0 + \ldots + s_0\sigma_{l-1},$$
$$s_l\sigma_0 + \ldots + s_1\sigma_{l-1}, \ldots, s_h\sigma_{l-2} + s_{h-1}\sigma_{l-1}, s_h\sigma_{l-1}). \tag{4.3}$$

If the select columns are deleted, and the resulting square determinant Δ expanded, we see that it takes the form of a sum of $l \times l$ determinants with columns the σ_r $(r = 0, 1, \ldots, -1)$. Since any one of these with two equal columns is zero, it follows that the only non-vanishing determinant among them is $|\sigma_0\sigma_1 \ldots \sigma_{l-1}|$. Therefore every surviving term has this determinant as a factor. However elementary reasoning shows that

$$|\sigma_0\sigma_1 \ldots \sigma_{l-1}| = \pm \prod_{\gamma_a \neq \gamma_b} (\gamma_a - \gamma_b) \neq 0. \tag{4.4}$$

The array (4.3) is the symbolic product of (4.4) with the array

$$\begin{array}{cccccccccccc} s_0 & s_1 & s_2 & . & . & . & s_h & 0 & 0 & . & . & . & 0 \\ 0 & s_0 & s_1 & s_2 & . & . & . & s_h & 0 & . & . & . & 0 \\ 0 & 0 & s_0 & . & . & . & & s_h & 0 & . & . & . & 0 \\ & & & & . & . & . & . & . & & & & \\ 0 & . & . & . & & & 0 & s_0 & s_1 & . & . & . & s_h \end{array} \tag{4.5}$$

of l rows and $h + l$ columns, according to the formal rules of determinant multiplication. We have therefore to show that the $l \times l$ determinant which remains when any h columns have been deleted from (4.5) is not zero.

This $l \times l$ determinant has the form of the representation of a Schur function $\{\lambda\}$ corresponding to a certain partition (λ) consisting of h numbers $\lambda_1, \ldots, \lambda_h$ arranged in decreasing order. For the theory of partitions and S-functions we refer to (**11**, chapters 5, 6). The partition (λ) is best defined

in this case by means of its conjugate partition (μ), which consists of l positive integers μ_i, not necessarily distinct, arranged in decreasing order. We set

$$\mu_i = h - d(i)$$

where $d(i)$ is the number of select, or assigned, derivatives of the sequence

$$u, u_x, u_{xx}, \ldots, u_x^{(n-1)}$$

which are encountered before the ith non-select derivative. From **(10**, chapter 6, 3.3**)** we see that the Schur function

$$\{\lambda\} = |s_{\mu_i - i + j}| \tag{4.6}$$

has the form of (4.5) after deletion of the select columns and transposition about the secondary diagonal.

On the other hand, from **(10**, chapter 6, 3.1**)** we have

$$|s_{\mu_i - i+j}| = \{\lambda\} = \frac{|c_i^{\lambda_j+h-j}|}{|c_i^{h-j}|}, \qquad (i, j = 1, \ldots, h), \tag{4.7}$$

where h is the number of select c_i's and i and j are indices of position in the determinants. We recall that the c_i are all distinct; the denominator in (4.7) is the Vandermonde's determinant which is equal to

$$\pm \prod_{j \neq i} (c_i - c_j),$$

and so is not zero. The numerator is a slightly more general type of determinant, which has been studied in **(1)** and shown to be different from zero. For this it is necessary that the c_i should be distinct and positive, and the powers $\lambda_j - j$ distinct, and these requirements are satisfied since the λ_j are non-increasing with j, while the c_i, being the select γ_i, are positive and distinct. A direct proof that the Schur function $\{\lambda\}$ is a symmetric polynomial of the c_i with non-negative coefficients has been given recently in **(8**, Theorem 1**)**. Thus (4.7) is different from zero in our case since the c_i are all positive. Combining this with (4.4) we see that the original subdeterminant of (4.1) is not zero, and this concludes the proof of the lemma.

The special case $h = 0$ is needed in connection with (2.6), and a sequence of applications with various values of h and l, one for each subgroup of Tables I and II, is needed in § 3 as stated there.

5. Verification of the solution. By **(3**, Theorem 3**)**, a piecewise analytic solution of (2.1), satisfying (2.2) and appropriate initial conditions, exists. Let the solution of (2.1) with given Cauchy data and boundary conditions (2.2), which is defined by the piecewise analytic expansions of **(3**, Theorem 3**)**, be constructed, and let us show that the solution u of (1.1) which we seek is actually the component v of the first equation of the first group of Table I. To show that v satisfies (1.1) we shall verify that the defining equations in

the left columns of Tables I and II hold, in succession, and use the last subgroup of equations of the first table. When the "definitions" are re-established the various boundary conditions will be automatically satisfied, in view of the equivalence (2.6) between the derivatives of u of a given order, and the variables $v_a \ldots$ of the same order. Thus it will be established that

$$u = v \tag{5.1}$$

is a solution of the original problem, since the algebraic verification of the initial conditions will be trivial. We shall use the uniqueness property of solutions of the first order system which are analytic on the closure of the sector domains.

Consider first Table I, and let us verify the relations in the left-hand column subgroup by subgroup. The first such relation, namely $v = u$, is taken as a hypothesis, or rather a definition of u. The second definition of the second subgroup is precisely the first differential equation and so is valid. To show that the first definition of the second subgroup holds, let us define

$$\xi_b = D_b u - v_b. \tag{5.2}$$

Then ξ_b is piecewise analytic on the closure of the sector domains R_i. Also

$$\begin{aligned} D_a \xi_b &= D_a D_b u - D_a v_b \\ &= D_b D_a u + C_{ab}[u] - v_{ab} - C_{ab}[v] \\ &= v_{ba} - v_{ab} + C_{ab}[u] - C_{ab}[v] \\ &= C_{ab}[\xi], \end{aligned} \tag{5.3}$$

where we have used the first three differential equations of Table I. With

$$C_{ab}[u] = D_a D_b u - D_b D_a u = \alpha D_a u + \beta D_b u,$$

where α and β are certain coefficients which we need not calculate explicitly, we have

$$C_{ab}[v] = \alpha v_a + \beta v_b$$

and therefore

$$C_{ab}[\xi] = \alpha \xi_a + \beta \xi_b.$$

Now in this case, $\xi_a = D_a u - v_a \equiv 0$. Thus ξ_b satisfies the homogeneous linear equation

$$D_a \xi_b = \beta \xi_b. \tag{5.4}$$

The initial conditions for ξ_b are also homogeneous, as follows from the definition of initial conditions for the v variables. If D_a is select, there is a homogeneous boundary condition for ξ_b on T. In this case only, discontinuities of the derivatives of ξ_b across G_a are in principle permitted, but the expansions of (3) applied to this equation show that all such jumps are here zero. Since G_a is the only characteristic surface of (5.4), it follows that ξ_b is analytic everywhere and so identically zero.

To verify the third subgroup of definitions, note that the last of these relations is now equivalent to the last differential equation of the preceding subgroup. Define

$$\begin{aligned} \xi_{cb} &= D_c D_b u - v_{cb} = D_c v_b - v_{cb} \,; \\ \xi_{ca} &= D_c D_a u - v_{ca} = D_c v_a - v_{ca} \,; \end{aligned} \tag{5.5}$$

then

$$\begin{aligned} D_a \xi_{cb} &= D_a D_c D_b u - D_c v_{cb} \\ &= D_c D_b D_a u + C_{acb}[u] - D_a v_{cb} \\ &= D_c v_{ba} + C_{acb}[u] - v_{cba} - C_{acb}[v] \\ &= C_{acb}[\xi], \end{aligned} \tag{5.6}$$

and likewise

$$D_b \xi_{ca} = C_{bca}[\xi] \,; \tag{5.7}$$

using the differential equations and previously established definitions. Here the $C_{acb}[\xi]$ expressions are linear homogeneous in the ξ variables with less than three indices: since all one-index ξ's are zero, (5.6) and (5.7) form a linear homogeneous system for ξ_{cb}, ξ_{ca}. Again, these functions satisfy homogeneous initial conditions. As above it follows in either case that ξ_{cb} and ξ_{ca} vanish identically.

The inductive procedure for the kth subgroup is similar: the last definition of the subgroup is true in view of the previously established definitions and the last differential equation of the preceding subgroup. We define $k - 1$ variables $\xi_{kh\ldots b}, \ldots, \xi_{kg\ldots a}$ as follows:

$$\begin{aligned} \xi_{kh\ldots b} &= D_k D_h \ldots D_b u - v_{kh\ldots b}, \\ \xi_{kg\ldots a} &= D_k D_g \ldots D_a u - v_{kg\ldots a}, \end{aligned} \tag{5.7}$$

there being a different D operator missing in each of these sequences of differentiations. Then

$$\begin{aligned} D_a \xi_{kh\ldots b} &= D_a D_k D_h \ldots D_b u - D\, v_{kh\ldots b} \\ &= D_k D_h \ldots D_b D_a u + C_{akh\ldots b}[u] - D_a v_{kh\ldots b} \\ &= D_k v_{hg\ldots ba} + C_{akh\ldots b}[u] - v_{kh\ldots ba} - C_{akh\ldots b}[v] \\ &= C_{akh\ldots b}[\xi] \,; \end{aligned} \tag{5.8}$$

and there are $k - 2$ similar equations of which the last is

$$D_h \xi_{kg\ldots a} = C_{hkg\ldots a}[\xi]. \tag{5.9}$$

Here the $C_{akh\ldots b}[\xi]$ are linear expressions containing the $\xi_{kh\ldots b}, \ldots, \xi_{kg\ldots a}$, as well as those of lower order (which are now known to be zero). Thus (5.7), ..., (5.8) form a self-contained linear homogeneous system, with homogeneous initial and boundary conditions. Since the $\xi_{kh\ldots b}, \ldots, \xi_{kg\ldots a}$ are analytic on the closure of each sector R_i they are identically zero, in view of the uniqueness theorem in **(3)**.

The proof that the last subgroup of defining relations holds for the solutions of the first order system is similar to the earlier steps of the induction. The only difference is that the operator $L_1[v]$ replaces the variable $v_{khg\ldots a}$ in the general step. Since this quantity does not appear in the final form (5.8) and (5.9) of the equations for the ξ's, this change has no effect on the result. This shows, then, that all defining relations of Table I are valid.

Let us show that the defining equations of Table II hold for each index group $(i)_q = (i_1 \ldots i_q)$ by induction on q for each of these groups. Let $(i)_{q-1} = (i_1 \ldots i_{q-1})$ and let us assume that the result has been proved for the $(i)_{q-1}$ group. First define

$$\xi_{(i)_q} = \frac{\partial^q u}{\partial x^{i_1} \ldots \partial x^{i_q}} - v_{(i)_q} = \frac{\partial}{\partial x^{i_q}} v_{(i)_{q-1}} - v_{(i)_q}. \tag{5.10}$$

We see that

$$\begin{aligned} D_a \xi_{(i)_q} &= D_a \frac{\partial}{\partial x^{i_q}} v_{(i)_{q-1}} - D_a v_{(i)_q} \\ &= D_a \frac{\partial}{\partial x^{i_q}} v_{(i)_{q-1}} - \frac{\partial}{\partial x^{i_q}} v_{a(i)_{q-1}} - C_{a i_1 \ldots i_q}[v], \end{aligned} \tag{5.11}$$

by the first differential equation of the group $(i)_q$. However the right side of (5.11) contains only v variables of q-order $(i)_{q-1}$ or less, and in view of the definition (2.5) of the commutator operator, this side of (5.11) will be zero, since all variables

$$v_{(i)_{q-1}}$$

have been proved equal to the corresponding partial derivatives of u. It follows by differentiation of the initial and boundary conditions that

$$\xi_{(i)_q}$$

vanishes identically.

For the typical kth subgroup of Table II, we have a system of k equations involving the variables

$$\begin{aligned} \xi_{kh\ldots b(i)_q} &= \frac{\partial}{\partial x^{i_q}} v_{kh\ldots b(i)_{q-1}} - v_{kh\ldots b(i)_q}, \\ \xi_{kg\ldots a(i)_q} &= \frac{\partial}{\partial x^{i_q}} v_{kg\ldots a(i)_{q-1}} - v_{kg\ldots a(i)_q}. \end{aligned} \tag{5.12}$$

Then, for instance, from the first differential equation of the subgroup, we have

$$\begin{aligned} D_a \xi_{kh\ldots b(i)_q} &= D_a \frac{\partial}{\partial x^{i_q}} v_{kh\ldots b(i)_{q-1}} - D_a v_{kh\ldots b(i)_q} \\ &= D_a \frac{\partial}{\partial x^{i_q}} v_{kh\ldots b(i)_{q-1}} - \frac{\partial}{\partial x^{i_q}} v_{kh\ldots a(i)_{q-1}} - C_{a i_q kh\ldots b(i)_{q-1}}[v]. \end{aligned} \tag{5.13}$$

Now the right side contains a commutator which includes only v variables of order $q - 1$ or less with respect to the $x^{i\nu}$, or else of order q but from a preceding subgroup of the qth group. As the identification of all these with the corresponding derivatives of u has already been made, we see by the definition (2.5) that the right side of (5.13) reduces to zero. Similarly, all other right sides, obtained by differentiation of the quantities in (5.12) by appropriate $D_{\flat} \ldots D_h$ operators, are seen to vanish. Homogeneous auxiliary conditions are applicable as before, and it follows that the variables ξ in (5.12) vanish identically.

Proceeding thus by induction we make all the identifications of the various q groups, and so identify all derivatives of $u = v$ with the appropriate v variables as foreshadowed in (2.3). This proves that u satisfies all the initial and boundary conditions. It remains now to show that u satisfies the original linear partial differential equation of order m. However this follows at once from the very last differential equation of Table I and the definition (1.9) of the operators $L_1(u)$ and $L_1[v]$. The existence proof is therefore complete.

THEOREM 1. *Let $L(u) = 0$ be an analytic linear differential equation of order m which is regularly hyperbolic with respect to analytic initial and boundary surfaces: $S{:}t = 0$ and $T{:}x = 0$. Let k_0 characteristic surfaces G_i issuing from $C = S \cap T$ into the region R be selected, and let k_0 of the quantities*

$$u, u_x, \ldots, u_x^{(m-1)}$$

be assigned on T in addition to Cauchy data on S. Then there exists a piecewise analytic solution u assuming the given initial and boundary values, and analytic except across the G_i where it is of class C^{m-1}.

This analytic solution is piecewise analytic in the strong sense described in (**2**, § 10); that is, it is analytic on the closures of the distinct sector domains D_h which separate the select characteristic surfaces. The solution must still be proved unique even within this class of functions, since there are many ways of setting up the corresponding first order system, and it is necessary to show that these distinct ways all lead to the same discontinuities of derivatives across the select characteristic surfaces.

To prove this, let us suppose that u is piecewise analytic in the above strong sense, that u is C^{m-1}, and analytic except across the select G_i; that $L(u) = 0$ and that homogeneous Cauchy data and boundary conditions have been assigned. Then u and its derivatives satisfy the system of Tables I and II, which is arranged according to any one of these alternative ways. However the data entering this system are all zero. Since a solution of the system is unique in the strongly piecewise analytic function class, (**3**) the solution of the system is identically zero. Hence $u \equiv 0$ as was to be proved.

As a further corollary we add that if every characteristic surface G_i issuing into the domain is select, then our solution is unique in the class of C^m functions. This now follows at once from the uniqueness in the C^1 class of solutions

of the first order system, when every characteristic root is real and not zero, and when all positive roots are select **(3**, § 10**)**.

We have seen in § 3 that lower order derivatives with respect to x can be permitted to appear in the boundary conditions. It is also true that lower order derivatives with respect to the other $N - 1$ variables can be accommodated in the same way. This is possible since we could establish the definitions required in working back to the mth order equation in a lexicographic sequence which takes account of first, order of the derivative, second, index (i) of the group, and third, ordering of the D_a operators within each subgroup. However it is not possible to permit oblique derivatives of an order equal to the highest order which occurs in the boundary condition, as can be seen even for hyperbolic equations of first or second order **(2)**. Such conditions will lead to inconsistencies if directional derivatives involved have characteristic directions.

We comment on the fact that non-analytic "kinks" can be chosen to occur on some, but perhaps not all, characteristic surfaces issuing into the region. Each such characteristic surface may be thought of as corresponding to a particular kind of wave, generated at the boundary. For a vibrating beam there will be flexural and shear waves, travelling at different speeds, for example. Our theorem shows that there is a solution, satisfying one boundary condition, in which only one type of wave is generated. With a suitable boundary condition, it is quite possible to have a solution in which only some other type of wave arises. In a physical problem, the appropriate linear combination of these solutions would have to be selected by some further conditions; usually these would be additional boundary conditions.

6. The non-analytic case. An existence theorem for hyperbolic equations of order m has been given by Leray **(9)** under an assumption of finite differentiability, and by means of analytic approximations for which uniform estimates are obtained through the use of energy integrals. More recently Gårding **(5)** has given a direct existence proof using only the energy integrals. These calculations refer to the Cauchy problem, and we shall here investigate their application to mixed problems as in the preceding sections. For second order equations, this aspect of mixed problems has been treated, for example in **(2, 7, 10)**.

The results which we obtain do generalize the known theorems for second order equations in two different ways, which will constitute Cases I and II below. However it has not been possible to attain the generality of the theorem for analytic equations, and a considerable gap remains to be filled. It should be remarked that for the case of two variables a thorough treatment by Picard's method has been given by Campbell and Robinson **(1)**, covering semilinear equations as well. The energy integral method has been applied to the linear problem in two variables by Thomée **(12)**.

In contrast to the analytic case, we must now assume that the differential equation (1.1) is regularly hyperbolic in the sense of Leray: that is, in effect,

that there exist timelike directions and that the normal cone is real and has no multiple points except the origin. This criterion will be fulfilled if the initial surface S is so situated that the equation is regularly hyperbolic (in the sense of § 1) with respect to S and to every surface T meeting S in a smooth hypercurve C.

When applied to a mixed problem, the energy integral formulae are modified by the presence of a boundary integral taken over the surface T. To complete the estimates we must show that this boundary integral form is semi-bounded. We therefore begin with an algebraic study of this boundary term, and will use the elegant notation of Hörmander **(6)** for the algebra of energy integrals.

Let the terms of highest order m in (1.1) be written

$$P(D)u, \qquad D_j = \frac{1}{i}\frac{\partial}{\partial x^j}, \tag{6.1}$$

where $P(D)$ is a polynomial of order m; and let $Q(D)$ be a real polynomial operator of order $m - 1$ in D. Then the quadratic form

$$F(D, \bar{D})\, u\, \bar{u} \equiv P(D)\, u\, \bar{Q}(\bar{D})\, \bar{u} - \bar{P}(\bar{D})\, \bar{u} \,.\, Q(D)u \tag{6.2}$$

is a divergence expression

$$-\, i \sum_j \partial/\partial x^j (G_j(D, \bar{D})\, u\, \bar{u}), \tag{6.3}$$

where the operators $G_i(D, \bar{D})$ are related to $F(D, \bar{D})$ by the equation

$$F(\zeta, \bar{\zeta}) = \sum_j (\zeta_j - \bar{\zeta}_j)\, G_j(\zeta, \bar{\zeta}). \tag{6.4}$$

Here $\zeta_j = \xi_j + i\eta_j$ and $\bar{\zeta}_j = \xi_j - i\eta_j$ are complex variables dual to D_i. In forming (6.3) we shall assume that the coefficients of $P(D)$ and $Q(D)$ are constants; this assumption can later be relaxed.

Writing the differential equation (1.1) in the form

$$L_u = P(D)u + B(D)u = f(x^i), \tag{6.5}$$

where $B(D)$ is an operator of order less than m, we integrate the expression $2\, Re\, Lu\, \bar{Q}(\bar{D})\bar{u}$ over a lens-shaped region R such as is described in **(3**, Figure 2**)**. This region is bounded by initial and final surfaces S_0 and S_t (t = const), and by a portion T_t ($x = 0$) of the boundary surface T. We find, on the one hand

$$i \int_R F(D\, \bar{D})\, u\, \bar{u}\, d\, V = \int_R Q(D, \bar{D}, u, \bar{u}, f)\, d\, V, \tag{6.6}$$

where the quadratic form $Q(D, \bar{D}, u, \bar{u}, f)$ is of order $m - 1$ or less in the D_i, and contains factors linear in f. On the other hand, by (6.3) we have

$$i \int_R F(D\, \bar{D})\, u\, \bar{u}\, d\, V = \int_{S_t - S_0} G_t(D\, D)\, u\, \bar{u}\, d\, S_t - \int_{T_t} G_x(D\, \bar{D})\, u\, \bar{u}\, d\, S_x, \tag{6.7}$$

the minus sign in the last integral being due to the convention that x shall be measured as increasing along the interior normal to T_t. Comparing (6.6) and (6.7), we find

$$(6.8)\quad \int_{S_t} G_t(D\bar{D})\, u\bar{u}\, d\,S_t - \int_{T_t} G_x(D\bar{D})\, u\bar{u}\, d\,S_x$$
$$= \int_{S_0} G_t(D\bar{D})\, u\bar{u}\, d\,S_t + \int_{R_t} Q(D, \bar{D}, u, u, f)\, d\,V.$$

The method of Leray and Gårding is based on the fact that if the auxiliary operator $Q(D)$ is so chosen that the sheets of its normal cone separate the sheets of the normal cone of $P(D)$, then the integral over S_t is positive definite (**4**, Lemma 3.1). For this purpose we may assume that the coefficients of $i^m D_t^m$ in $P(D)$, and of $i^{m-1} D_t^{m-1}$ in $Q(D)$, are both $+ 1$.

Now let us suppose that $P(D)$ and $Q(D)$ have variable but sufficiently often differentiable coefficients. Then (6.3) is modified by the addition of derivatives of order lower than $m - 1$, and a quadratic form in such derivatives of u, $\bar{u}$ will appear in (6.7). These terms may be absorbed in the integral over R_t in (6.8), which is thus not changed in form. Also, in this case, the integral over S_t can be made positive definite in all derivatives of u of all orders $< m - 1$, by the addition of derivatives to the integrand $G_t(D\,D)\, u\,\bar{u}$ of orders not greater than $m - 2$. Again, such terms in the new integrand $G_t(D\,\bar{D})\, u\,\bar{u}$ over S_t can be counterbalanced by terms in the volume integral containing derivatives of order no higher than $m - 1$. Consequently (6.8) holds unchanged in form for the case of variable coefficients as well, with a somewhat different quadratic form $Q(D, \bar{D}, u, \bar{u}, f)$ in the volume integral. Set

$$(6.9)\qquad E_k(t) = \int_{S_t} \sum_{0 \leqslant \alpha \leqslant k} |D^\alpha u|^2\, d\,S_t,$$

the summation being taken over all essentially distinct partial derivatives of u of order less than $k + 1$. By the positive definiteness of the integrand in the integral over S_t, we find, (**4**, Theorem 2.1; **5**, Theorem 3.1) that there exists a constant $c > 0$, depending only on the differential equation and the domain, such that

$$(6.10)\qquad \int_{S_t} G_t(D, \bar{D})\, u\,\bar{u}\, d\,S_t > c^{-1} E_{m-1}(t) - c\, E_{m-2}(t),$$

for every t and all $u \in C^{m-1}$.

We may express $E_{m-2}(t)$ as the integral of the time-derivative of its integrand: this leads to an estimate

$$(6.11)\qquad E_{m-2}(t) < E_{m-2}(0) + K \int_0^t E_{m-1}(t)\,dt.$$

It now follows from (6.8) and (6.10) that

$$
\begin{aligned}
c^{-1} E_{m-1}(t) &< c\, E_{m-2}(t) + \int_{S_t} G_t(D, \bar{D})\, u\, \bar{u}\, d\, S_t \\
&< c\, E_{m-2}(0) + c\, K \int_0^t E_{m-1}(\tau) d\tau \\
&\quad + \int_{S_0} G_t(D, \bar{D})\, u\, \bar{u}\, d\, S_t + \int_{R_t} Q(D\, \bar{D}\, u\, \bar{u}\, f) dV \\
&\quad + \int_{T_t} G_x(D\, \bar{D})\, u\, \bar{u}\, dS_x.
\end{aligned} \tag{6.12}
$$

By Schwarzian estimations of the third and fourth terms on the right-hand side of this last inequality, we find

$$
\begin{aligned}
E_{m-1}(t) &< c^2 E_{m-2}(0) + K_1 E_{m-1}(0) + K_2 \int_0^t E_{m-1}(\tau) d\tau + I_B(t) \\
&< K_0 + K_2 \int_0^t E_{m-1}(\tau)\, d\, D + I_B(t)
\end{aligned} \tag{6.13}
$$

where we have written

$$
I_B(t) = \int_{T_t} G_x(D, \bar{D})\, u\, \bar{u}\, d\, S_x. \tag{6.14}
$$

Here K_0 and K_2 are constants depending on all the data of the problem, but not on u.

Now let us suppose that we can prescribe a similar estimate

$$
I_B(t) < K_4 + K_5 \int_0^t E_{m-1}(\tau) d\tau, \tag{6.15}
$$

for the boundary integral. By standard methods (9; 4, Lemma 1.2) we can now establish a conventional L^2 estimate

$$
E_{m-1}(t) < (K_0 + K_4) \exp\, [(K_2 + K_5)t]. \tag{6.16}
$$

Integration with respect to t leads to L^2 estimates over the entire domain R_t, and the process of solution, using analytic approximation together with Sobolev's lemma and Ascoli's selection theorem, then proceeds as in (9, p. 162). Further details will not be presented here.

To summarize this discussion, we state

LEMMA 2. *Let initial and boundary surfaces S and T subtend exactly k characteristic surfaces of the regularly hyperbolic equation*

$$
Lu = f, \tag{6.17}
$$

the surfaces and functions present being of class $[\frac{1}{2}N] + h + m$ in the closure of the region R. Let zero Cauchy data be assigned on S, and let k of the derivatives $u, u_x, \ldots, u_x^{m-1}$ be assigned the value zero on T. Then if the boundary integral $I_B(t)$ satisfies an estimate (6.15), *there exists a solution $u \in C^{h+m}$ of* (6.17) *which satisfies these conditions.*

We note in passing that the problem with non-homogeneous boundary conditions can be reduced to the form above by subtraction of a function which satisfies the initial and boundary conditions.

In Cases I and II below we shall need quite different methods to establish the inequality (6.15). The work falls into two parts, namely, an analysis of the case of constant coefficients, and an adaptation of this case to the more general situation with variable coefficients.

7. The boundary form. We consider first the particular case of constant coefficients in the differential operator, and use Fourier transforms to estimate the integral over T_t. Let us denote by $\tilde{u} = \tilde{u}(\xi_i)$ the Fourier transform

$$\tilde{u}(\xi_i) = \int e^{2\pi i(\xi, x)} u(x^\rho, t)\, dx^\rho\, dt$$

of a function of x_ρ and t, defined as equal to $u(x^\rho, t)$ on T_t and zero elsewhere. Here also

$$(\xi, x) = \xi_0 t + \sum_\rho \xi_\rho x_\rho.$$

The inverse transform is easily written down, and we note that ξ_ρ is now dual to $D_\rho = -i\partial/\partial x_\rho$. Thus, we shall need to distinguish differentiation with respect to the transverse variable x (across T), by writing

$$G_x(D, \bar{D}) = G\,(D_t, D_x, \bar{D}_t, \bar{D}_x).$$

Now Parseval's theorem **(6)** shows that

$$\text{(7.1)} \quad I_B(t) = \int_{T_t} G_x(D_t, D_x, \bar{D}_t, \bar{D}_x)\, u\, \bar{u}\, d\, S_t = \int G_x(\xi_t, D_x, \xi_t, \bar{D}_x)\, \tilde{u}\, \bar{\tilde{u}}\, dS_t.$$

This last integrand is a quadratic form in the variables $D_x{}^j\, \tilde{u}$, ($j = 0, 1, \ldots, m-1$), which are independently defined when regarded as functions of the ξ_t. Thus in the case of constant coefficients we are led to study the algebraic properties of this quadratic form. We single out the variable $\zeta_x = \xi_x + i\eta_x$ and set all other variables ζ_t in G_x equal to their real parts ξ_t. From (6.4) it is found that

$$\text{(7.1)} \qquad G_x(\xi_t, \zeta_x, \xi_t, \bar{\zeta}_x) = \frac{F(\xi_t, \zeta_x, \xi_t, \bar{\zeta}_x)}{\zeta_x - \bar{\zeta}_x}.$$

Let us now write

$$\text{(7.2)} \qquad \begin{aligned} P(\zeta) &= P(\zeta_x) = P(\zeta_x, \xi) = \sum_{r=0}^{m} a_r(\xi_t)\zeta^r, \\ Q(\zeta) &= Q(\zeta_x) = Q(\zeta_x, \xi) = \sum_{s=0}^{m-1} b_s(\xi_t)\zeta^s. \end{aligned}$$

Then, dropping mention of the ξ_t variables ($i \neq x$), we find

$$G_x(\zeta_x, \bar{\xi}_x) = \frac{P(\zeta_x)Q(\bar{\xi}_x) - P(\bar{\xi}_x)Q(\zeta_x)}{(\zeta_x - \bar{\xi}_x)}$$

$$= \frac{\sum_{r=0}^{m} a_r \zeta_x^r \sum_{s=0}^{m-1} b_s \bar{\xi}_x^s - \sum_{r=0}^{m} a_r \bar{\xi}_x^r \sum_{s=0}^{m-1} b_s \zeta_x^s}{\zeta_x - \bar{\xi}_x}$$

$$= \sum_{r=0}^{m} \sum_{s=0}^{m-1} a_r b_s \left(\frac{\bar{\xi}_x^s \zeta_x^r - \zeta_x^s \bar{\xi}_x^r}{\zeta_x - \bar{\xi}_x} \right).$$

Since the expression in parentheses is the sum of the geometric series

$$\pm \sum_{k=0}^{|r-s|-1} \bar{\xi}^k \zeta^{|r-s|-1-k},$$

(the $+$ sign being taken if $r > s$, the minus if $r < s$), we find, after some rearrangement,

$$G_x(\zeta_x, \bar{\xi}_x) = \sum_{p,q=0}^{m-1} c_{pq} \zeta_x^p \bar{\xi}_x^q, \tag{7.3}$$

$$c_{pq} = \sum_{s=0}^{\min(p,q)} (b_s a_{p+q+1-s} - a_s b_{p+q+1-s}).$$

Here a_s is taken as zero for $s > m$, while $b_s = 0$ for $s > m - 1$. It may be noted that $c_{pq} = c_{pq}(\xi_t)$ is homogeneous of degree $2m - 2 - p - q$ in the variables ξ_t.

When k homogeneous boundary conditions are assigned, then in effect k rows and columns of the coefficient matrix c_{pq} are deleted, since the corresponding terms fall out. The estimate we require is essentially that the remaining, or residual, quadratic form, be non-positive. We consider two cases here: when it is negative definite, and when it is zero.

Let the residual form be negative definite; then by altering the k deleted rows and columns we can arrange that the new (enlarged) form should also be negative definite. Then an estimate of the type of (**5**, Lemma 3.1) will hold, though in the opposite direction. This property of the constant-coefficients case, which is a local property for variable coefficients, enables us to deduce the analogue of (**5**, Theorem 3.1) which applies to the case of variable coefficients. It now reads

$$\begin{aligned} I_B(t) &= \int_{T_t} G_x(D_t, D_x, \bar{D}_t, \bar{D}_x)\, u\, \bar{u}\, d\, S_t \\ &< -c^{-1}E_{m-1}(t) + cE_{m-2}(t), \qquad c > 0, \end{aligned} \tag{7.4}$$

where

$$E_j(t) = \int_{T_t} \sum_{|\alpha| < j} |D^\alpha u|^2\, d\, S_t. \tag{7.5}$$

We must remember that the assigned homogeneous boundary data are inserted on the left in (7.4). The first term on the right in (7.4) can be replaced

by zero; we shall now estimate the second one, in much the same way as in (6.11). Let each point of T_t be joined to a point of S_0 by a line $x + t =$ const, $x_\rho =$ const, $\rho = 1, \ldots, N-2$; and let us express the integrand of $\tilde{E}_{m-2}(t)$ as the integral of its derivative along this line. Thus derivatives of order $m-1$ or less appear; and integration over T_t leads to integrals over a "triangular" portion of R_t making their appearance. Application of Schwarz' inequality now leads at once to the estimate (6.15) which we require.

This method of negative definite character for the residual quadratic form will be used in Case I below.

For Case II, where the residual quadratic form vanishes identically at every point of T, we must use a different approach to gain the result for variable coefficients. Let us use the fact that the coefficients are continuous, and, given a fixed function u, together with an arbitrary positive number ϵ, subdivide the boundary surface T_t into a finite number of portions T_h, in each of which the oscillation of the coefficients is less than ϵ. Select a point x_0^h in each T_h, and write

$$\begin{aligned}\int_{T_h} G_x(D_t D_x \bar{D}_t \bar{D}_x)\, u\bar{u}\, d\, S_T &\\ = \int_{T_h} G_{x_0}(D_t D_x \bar{D}_t \bar{D}_x)\, u\bar{u}\, d\, S_t &\\ + \int_{T_h} R(D_t, D_x, \bar{D}_t \bar{D}_x)\, u\bar{u}\, d\, S_T,&\end{aligned}$$

where

$$G_{x_0}$$

is the boundary form with constant coefficients evaluated at x_0, and the coefficients in the remainder term $R(D, \bar{D})u\bar{u}$ are all less than ϵ in magnitude.

Suppose now that u satisfies the homogeneous boundary conditions; then by hypothesis the first integral on the right vanishes. The second integral can be estimated to be less than

$$\epsilon\, m^2 \int_{T_h} \sum_{|\alpha| \le m-1} |D^\alpha u|^2\, d\, S_T.$$

It follows by summation over the T_h that the boundary integral $I_B(t)$ is less than

$$\epsilon\, m^2 \tilde{E}_{m-1}(t)$$

in magnitude. However, u and therefore $\tilde{E}_{m-1}(t)$, are fixed, and ϵ is arbitrary. Consequently $I_B(t)$ must vanish. For the variable coefficient problem, it is therefore sufficient that the coefficients at each point of T_t should lead to a vanishing residual matrix. We employ this result in Case II below.

8. Case I: $k = m - 1$. The corner element $c_{m-1,m-1}$ of the coefficient matrix of the above quadratic form is a polynomial of degree zero in the

ξ_t; it is in fact $a_m b_{m-1}$. If this element is negative then the conditions of the lemma will be fulfilled when the function and its first $m-2$ derivatives $u, u_x, \ldots, u_x^{(m-2)}$ are given as zero on T. Now this boundary condition will be appropriate if $k = m - 1$ characteristic surfaces lie between S and T: we assume this. Consequently one characteristic surface lies between T and the portion of S prolonged beyond the edge $C = S \cap T$. Define an auxiliary co-ordinate $z = t\cos\alpha + x\sin\alpha$, and let α range from $\alpha = 0$ to $\alpha = \frac{1}{2}\pi$. The coefficient of D_z^m in $P(D)$ is equal to 1 when $\alpha = 0$ and z coincides with t. Since this coefficient vanishes when the surface $z = \text{const}$ is characteristic, and changes sign at a simple characteristic surface, it vanishes once for $0 < \alpha < \frac{1}{2}\pi$ and is therefore negative. For $\alpha = \frac{1}{2}\pi$ it is a_m which is thus negative.

To make b_{m-1} positive, we should, in view of this discussion and of the fact that the characteristic surfaces of $Q(D)$ must separate those of $P(D)$, arrange that all $m-1$ of these surfaces should lie between S and T. That is, T must be spacelike with respect to $Q(D)$, or equivalently the normal to T must be timelike. We shall *assume* that it is possible to find an auxiliary operator which has this property. For example, if $m = 2$, the order of $Q(D)$ is 1 and the characteristic surface can be chosen to have any direction between S and T. In the general case, it will be possible to find such an operator whenever T lies sufficiently close to the single characteristic surface G which lies outside the domain between S and T.

Together with Lemma 2 this demonstrates the following.

THEOREM 2. *Let $k = m - 1$ and suppose there exists an operator $Q(D)$ separating the sheets of $P(D)$ such that all $m - 1$ characteristic surfaces of $Q(D)$ lie between S and T in the region R. Then there exists a solution of $Lu = f$, with given Cauchy data on S, and with given values for the $m - 1$ quantities u, u_x, $u_x^{(2)}, \ldots, u_x^{(m-2)}$ on T.*

When the coefficients of the differential equation are independent of x, it is possible to show that two other sets of $m - 1$ boundary conditions can be reduced to the set just treated. We write $Lu = u_x^{(m)} + \alpha_{m-1}u_x^{(m-1)} + \ldots + \alpha_0 u$, where α_{m-k} is a differential operator of order k in D_t, D_y, with coefficients independent of x.

COROLLARY. *Let the coefficients in* (1.1) *be independent of x. Then there exists a solution of the preceding mixed problem when the boundary conditions are*

(a) $$u_x^{(h)} = 0, \qquad h = 0, 1, \ldots, m-3, m-1, \qquad h \neq m-2$$

or

(b) $$u_x^{(h)} = 0, \qquad h = 1, 2, \ldots, m-1, \qquad h \neq 0.$$

To prove (a) let us suppose that v is a solution of this problem in the analytic case, and let us show that $v = u_x + \alpha_{m-1}u$, where u is a solution of a suitably selected problem with $u, u_x, \ldots, u_x^{(m-2)}$ vanishing on T. Since

all $m - 1$ characteristic surfaces between S and T are select, it follows from the reduction of Theorem 1 and the uniqueness theorem of (3, § 9) that the solution is unique. Now let u satisfy $Lu = g$, where g is a solution, vanishing on T, of the first order linear partial equation

$$\frac{\partial g}{\partial x} + \alpha_{m-1}\, g = f.$$

Since the coefficient of $\partial g/\partial x$ is not zero, and since f is supposed analytic, such an analytic solution g exists and is uniquely determined. Since $g = \int f\, ds$, where the integration is taken along a characteristic curve, we can find L^2 estimates for g and its derivatives if such estimates are given for f.

Formal calculation, using the non-dependence on x of the coefficients, now shows that the combination $w = u_x + \alpha_{m-1}u$ satisfies $w = 0$, $w_x = 0, \ldots,$ $w_x^{(m-3)} = 0$, on T, while $w_x^{(m-1)} = u_x^{(m)} + \alpha_{m-2}u_x = \alpha_{m-2}ux^{(m-2)} - \ldots$ $- \alpha_0 u + g$, which latter expression also vanishes on T by the boundary conditions for u and g.

Now

$$Lw = L(u_x + \alpha_{m-1}\, u) = \left(\frac{\partial}{\partial x} + \alpha_{m-1}\right) Lu = \left(\frac{\partial}{\partial x} + \alpha_{m-1}\right) g = f,$$

so that w is an analytic solution of the case (a). Hence, by the uniqueness property, $v = w = u_x + \alpha_{m-1}u$. Since we have found L^2 bounds for u and its derivatives, it now follows that such bounds can be obtained for v if one degree of differentiability extra is assumed for the non-analytic problem. The remainder of the existence proof now follows the conventional methods and so is omitted.

The demonstration of case (b) is similar in principle, but a different device is used. We note that the "coefficient" α_0, which is a differential operator of order m in the other $N - 1$ derivatives, contains just those terms of Lu not involving d/dx, and so is regularly hyperbolic in the $N - 1$ variables. Also, the edge $C = S \cap T$ is a spacelike surface relative to α_0, so that the Cauchy problem $\alpha_0 z = 0$, with Cauchy data on C, is a correctly set problem within the boundary T. We will show that a solution v of case (b) is equal to a combination

$$w = u_x^{(m-1)} + \alpha_{m-1}u_x^{(m-2)} + \ldots + \alpha u,$$

when certain preliminary reductions have been made. As the non-homogeneous boundary conditions corresponding to Theorem 2 can be set up by a substitution, we shall consider the problem $v_x = f_1$, $v_x^{(2)} = f_2, \ldots v_x^{(m-1)} f_{m-1}$ on T, with $Lv = 0$ in R, and, as usual, zero Cauchy data. Subtracting from this a suitable solution of the problem with $u_x^{(m-1)}$ not given on T, we can suppose without loss of generality that $f_1 = f_2 = \ldots = f_{m-2} = 0$.

Now let u be that solution of $Lu = 0$ with $u = u_x = \ldots = u_x^{(m-3)} = 0$ on T, with $u_x^{(m-2)} = z$ on T, where z is the solution of $\alpha_0 z = -f_{m-1}$ on T.

Straightforward calculation shows that w, defined above, satisfies $Lw = 0$ with $w_x = w_x^{(2)} = \ldots = w_x^{(m-2)} = 0$ on T, while

$$w_x^{(m-1)} = u_x^{(2m-2)} + \alpha_{m-1}u_x^{(2m-3)} + \ldots + \alpha_1 u_x^{(m-1)} = -\alpha_0 u_x^{(m-2)} = f_{m-1}$$

on T. Thus w is an analytic solution of the problem and so is equal to v. Hence estimates for u can be applied now to v, provided that $m - 1$ extra degrees of differentiability are assumed for the original problem. This completes the reduction of case (b) to the conventional energy integral method.

9. Case II: Symmetry with respect to T. The second circumstance in which it can be shown that the residual quadratic form can be bounded above is when the hyperplane $T : x = 0$ is a plane of symmetry for the characteristic cone of the hyperbolic differential operator. We shall here restrict consideration to equations of even order, as it is necessary, for the odd order case, to make rather lengthy changes in the "analytic" existence theorems to cover this situation.

Thus let Lu be a regularly hyperbolic operator of even order $m = 2l$, of which the highest order terms contain D_x^2 but not odd powers of D_x, at any rate for $x = 0$. Then T is a plane of symmetry as stated above. If now the terms of order $2l$ are written as in (7.2), it is seen that $a_r(\xi_i) = 0$ for r odd. Let us take $Q(\zeta) = \partial P/\partial\zeta_i$; as shown in (**8**, p. 140) the sheets of the cone of Q will separate those of P as required for the formulation of the estimates. Thus the odd terms $b_s(\xi_i)$ in $Q(\zeta)$ will likewise vanish.

From (7.3) it is seen that in

$$c_{pq} = \sum_{s=0}^{\min(p,q)} (b_s a_{p+q+1-s} - a_s b_{p+q+1-s})$$

the sum of indices of the a and b coefficients is $p + q - 1$, which is odd whenever $p + q$ is even. It follows that each term will contain a vanishing factor when $p + q$ is even, and therefore that $c_{pq} = 0$ for $p + q$ even. Thus about half the terms in the matrix are zero, including all diagonal terms, all terms twice removed from the diagonal, and so on.

From the symmetry of the characteristic surfaces relative to the boundary T, we see that half of the sheets lie in R, and thus l boundary conditions are appropriate.

THEOREM 3. *Let $Lu = f$ be a regularly hyperbolic equation of even order $2l$, such that the boundary surface T is a plane of symmetry relative to the characteristic cone at each point of T. Then there exists a solution of the $C^{2l+[\frac{1}{2}N]+1}$ mixed problem with zero data assigned on T for either*

(a) *l derivatives of even order: $u, u_x^{(2)}, \ldots u_x^{(2l-2)}$ or*

(b) *l derivatives of odd order: $u_x, u_x^{(3)}, \ldots u_x^{(2l-1)}$.*

Proof. In case (a), an element c_{pq} belongs to the residual matrix only if both p and q are odd, and thus $p + q$ is even. Hence the residual part of the

quadratic form is identically zero. Similarly, in case (b), the residual part contains elements with both p and q even so that c_{pq} vanishes and the quadratic form is zero. An application of Lemma 2 completes the proof.

For the case $m = 2l = 2$ we can use the Lorentz transformation to show that the symmetry requirement can always be satisfied.

10. Signature of the quadratic form. We conclude with some remarks on the algebraic structure of the quadratic form $G_z(\zeta_z, \bar{\zeta}_z)$. An adaptation of Gårding's analysis (**5**, Lemma 1.1) to the case where complex roots are present shows that the signature of this quadratic form is always compatible with the number of boundary conditions suggested by the arrangement of characteristic surfaces. However, the coefficients of this quadratic form are coefficients of the variables ξ_ρ dual to t and x^ρ, $(\rho = 1, \ldots, N-2)$, and consequently the eigenvectors corresponding to the negative eigenvalues of the coefficient matrix depend on the ξ_ρ. As these eigenvectors are linear combinations of the transforms of the x-derivatives, it follows that in general k linear conditions (with coefficients depending on the ξ_ρ) of the transforms $\tilde{u}_z^{(h)}(\xi_\rho)$ are required as boundary conditions in order that the residual matrix should be bounded above. Upon transformation back to the t, x^ρ variables, these relations would become integral conditions of convolution type on the derivatives of u. It seems probable that Gårding's direct method could be modified to include this rather unconventional type of condition.

11. A remark. I wish to correct a misstatement in the paper (**3**) on first order systems. On page 154, the fourth line from the bottom of the page should read "Let a non-singular analytic family of surfaces S_t fill R in such fashion that through each point of R there passes one and only one surface S_t of the family, and that $S_1 = S_{t-1}$."

12. Acknowledgments. I am indebted to W. P. Brown for assistance with the algebra of Lemma 1. Professors K. O. Friedrichs and A. Robinson provided criticism which has led to certain improvements and clarifications, and for which I am most grateful.

References

1. L. L. Campbell and A. Robinson, *Mixed problems for hyperbolic differential equations*, Proc. Lond. Math. Soc. (3), *5* (1955), 129–47.
2. G. F. D. Duff, *A mixed problem for normal hyperbolic linear partial differential equations of second order*, Can. J. Math., *9* (1957), 141–60.
3. ——— *Mixed problems for linear systems*, Can. J. Math., *10* (1958), 127–60.
4. L. Gårding, *Dirichlet's problem*, Math. Scand., *1* (1953), 55–72.
5. ——— *Solution directe du problème de Cauchy pour les équations hyperboliques*, Proc. Coll. Int. du C.N.R.S. LXXI (Paris, 1956), 71–90.
6. L. Hörmander, *On the theory of general partial differential operators*, Acta Math., *94* (1955), 161–248.
7. M. Kryzyanski and J. Schauder, *Quasi-lineare Differentialgleichungen zweiter Ordnung vom hyperbolischen Typus, Gemischte Randwertaufgeben*, Studia Math., *6* (1936), 152–89.
8. L. Kuipers and B. Meulenbeld, *Symmetric polynomials with non-negative coefficients*, Proc. Amer. Math. Soc., *6* (1955), 88–93.
9. J. Leray, *Hyperbolic differential equations* (Princeton, 1953).
10. J. L. Lions, *Problèmes aux limites en théorie des distributions*, Acta Math., *94* (1955), 13–153.
11. D. E. Littlewood, *The theory of group characters* (Cambridge, 1939).
12. V. Thomée, *Estimates of the Friedrichs-Lewy type for mixed problems in the theory of linear hyperbolic differential equations in two independent variables*, Math. Scand., *5* (1957), 93–113.

University of Toronto

AN IMPROVED RESULT CONCERNING SINGULAR MANIFOLDS OF DIFFERENCE POLYNOMIALS

RICHARD M. COHN

1. Introduction. Let $\mathfrak{K}$ be a difference field of characteristic 0, $\mathfrak{M}$ an irreducible manifold of effective order n over $\mathfrak{K}\{y\}$, and F an algebraically irreducible difference polynomial in $\mathfrak{K}\{y\}$ of effective order $n + k$, $k > 0$, which vanishes on $\mathfrak{M}$. In an earlier paper (**2**, p. 447) I gave necessary conditions, restated below as (a), (b), and (c) of the main theorem, for $\mathfrak{M}$ to be an essential singular manifold of F. These conditions are analogous to the low power criterion of Ritt (**1**, p. 65) for the corresponding problem of differential algebra. Like that criterion they depend, in the special case that $\mathfrak{M}$ is the manifold of y, only on which power products appear effectively in F. Unlike the low power criterion, however, conditions (a), (b), and (c) are only necessary, not sufficient. I have proved the following results (**2**, p. 459; **4**) concerning sufficiency:

(1) *if $k = 1$, the conditions are never satisfied, so that $\mathfrak{M}$ is not an essential singular manifold of F;*

(2) *if $k = 2$, $n = 0$, the condition is both necessary and sufficient;*

(3) *if $k > 2$, the condition is not sufficient, even if $n = 0$. Moreover, no condition dependent only on which power products of y and its transforms appear effectively in F is sufficient in the special case that $\mathfrak{M}$ is the manifold of y.*

I shall now show that the restriction $n = 0$ may be removed from (2). Hence, there is a close analogy with the situation in differential algebra described by the low power theorem in the case that the effective order of the difference polynomial F exceeds by 2 the effective order of the manifold $\mathfrak{M}$, but only a partial analogy in all other cases.

The proof for $k = 2$ and "general" n is based on a preparation theorem suggested by the preparation theorem used by Ritt for differential polynomials. The preparation theorem of difference algebra (restricted to the case that $\mathfrak{K}$ is inversive) consists of the relations (3) and (4) of § 5 between F and the first polynomial A of the characteristic set of the reflexive prime difference ideal with manifold $\mathfrak{M}$. The conditions (a), (b), and (c) of the main theorem are equivalent to the conditions (α) and (β), stated in § 8, for (3) and (4). These conditions in turn imply that $\mathfrak{M}$ is an essential singular manifold of F in the case $k = 2$. The proof of this is accomplished by a minor modification of the power series method used in (**4**) for the special case $n = 0$ and the conditions (a), (b), and (c).

Received July 11, 1958. This investigation was supported in part by a grant from the Rutgers University Research Fund.

2. The *weight function* $f(\theta)$ of a term

$$\sigma y^{a_0} y_1^{a_1} \dots y_r^{a_r}, \qquad \sigma \neq 0, \sigma \in \mathfrak{K},$$

of a difference polynomial of $\mathfrak{K}\{y\}$ is defined to be the polynomial $a_0 + a_1\theta + \dots + a_r\theta^r$. The indeterminate θ is called the *weight parameter*. The *weight* of this term for a value τ of the weight parameter is $f(\tau)$. If an element c whose transform is defined to be c^τ is substituted formally into a term, then the exponent of c in the result is the weight of the term for the value τ of the weight parameter.

Let F and $\mathfrak{M}$ be as in § 1. Denote by α a generic zero of $\mathfrak{M}$ and by Σ the reflexive prime difference ideal with manifold $\mathfrak{M}$. Let A be an algebraically irreducible polynomial in Σ of effective order n—if $\mathfrak{K}$ is inversive, A is the first polynomial of a characteristic set of Σ or one of its transforms. If P is any polynomial of $\mathfrak{K}\{y\}$, then $\bar{P}$ is to denote the polynomial of $\mathfrak{K} \langle\alpha\rangle \{z\}$ which is obtained from P by the substitution $y = z + \alpha$, and P^* the polynomial consisting of the terms of least degree of $\bar{P}$. We can now state the main theorem.

THEOREM. *In order that $\mathfrak{M}$ be an essential singular manifold of F it is necessary that:*

(a) *there exist a term of $\bar{F}$ which is of lower weight than any other term for every positive value $\tau < 1$ of the weight parameter,*

(b) *there exist a term of $\bar{F}$ which is of lower weight than any other term for every value $\tau > 1$ of the weight parameter,*

(c) *every solution of F^* be a solution of A^*.*

These conditions are sufficient if $k = 2$.

It only remains to prove sufficiency in the case $k = 2$. The rest of this paper is devoted mainly to this proof. In the last section a method is given for testing conditions (a), (b), and (c) constructively if a beginning of a characteristic sequence of the ideal Σ is known.

3. Proof of a lemma. The first lemma to be proved concerns polynomial rings, the second, difference rings.

LEMMA 1. *Let $\mathfrak{K}$ be a field of characteristic* 0, Π *a prime ideal in the polynomial ring $\mathfrak{R} = \mathfrak{K}[u_1, \dots, u_q; x_1, \dots, x_r]$, the u_i forming a parametric set for* Π. *Let $A_1, \dots, A_r$ be a characteristic set for* Π *with A_i introducing x_i. Let F be a polynomial of $\mathfrak{R}$. Then there exists a polynomial S of $\mathfrak{R}$ which is not in* Π *and an integer t such that*

$$SF = \sum_{i=1}^{t} L_i A_1^{p_{i,1}} A_2^{p_{i,2}} \dots A_r^{p_{i,r}}, \tag{1}$$

the L_i being polynomials of R which are not in Π, *and the $p_{i,j}$, $i = 1, \dots, t$; $j = 1, \dots, r$, constituting t distinct sets of non-negative integers.*

Proof. A polynomial of $\mathfrak{R}$ is said to be of *class* k if it effectively involves x_k but no x_i, $i > k$. The conclusion of the lemma follows trivially if F is free of all the x_i. We shall prove by induction on the class that it is valid for other polynomials in the following strengthened form: Let a_i denote the degree of A_i in x_i, $i = 1, \ldots, r$. Then if F is of class k and degree f in x_k it is possible to find a relation (1) in which S and the L_i are free of the x_i, $i > k$, and the power products in the A_i which occur on the right-hand side of (1) involve no A_i, $i > k$, and are of degree in A_k less than or equal to the greatest integer h_k not exceeding f/a_k.

The strengthened result holds if F is of class 1. For, if $h \geqslant 0$ is the greatest power of A_1 which divides F, we may write $F = LA_1{}^h$. This expression is of the form (1) and meets the added conditions.

Let F be of class $k > 1$, and assume the strengthened conclusion to have been proved for all polynomials of lower class. Let f be as before. If $f < a_k$, we use the expression

$$F = F_0 + F_1x_k + \ldots + F_fx_k^f,$$

each F_i being free of x_i, $i \geqslant k$. For each F_i, $0 \leqslant i \leqslant f$, we find an expression of the form of (1):

$$S_iF_i = \sum_{j=1}^{t_i} L_{ij}A_1^{p_{i,j,1}} \ldots A_{k-1}^{p_{i,j,k-1}}, \tag{2}$$

where S_i and the L_{ij} are not in Π and are of class less than k.

Let S be the product of the S_i. Substituting from the expressions (2) into

$$SF = (S/S_0)S_0F_0 + (S/S_1)S_1F_1x_k + \ldots + (S/S_f)S_fx_k^f$$

and combining terms involving equal power products of the A_i we obtain an expression of the form (1) for SF. The coefficients L_i of this expression are polynomials in x_k of degree less than a_k, with coefficients free of x_i, $i \geqslant k$, and not in Π. Hence, the L_i themselves are not in Π. Clearly, S is not in Π, and A_k does not appear in the power products; so that the strengthened conclusion is valid for F.

We now suppose that $f \geqslant a_k$, and that the strengthened conclusion has been demonstrated for all polynomials of class k and degree less than f in x_k. Applying the division algorithm to F and $A_k^{h_k}$ we find a relation

$$JF - MA_k^{h_k} = R,$$

where J, M, and R are polynomials free of x_i, $i > k$, J is not in Π, M is of degree less than f in x_k, and R is of degree less than h_ka_k in x_k. By the assumption made at the beginning of this paragraph there exists a polynomial S_1 not in Π and free of x_i, $i > k$, such that S_1R is a linear combination of power products of $A_1, \ldots, A_k$ of degree less than h_k in A_k, the coefficients of these power products being polynomials not in Π and free of x_i, $i > k$. By the case $f < a_k$ previously disposed of there exists a polynomial S_2 not in Π and free of x_i, $i > k$, such that S_2M is a linear combination of power products of

$A_1, \ldots, A_{k-1}$, the coefficients of these power products being polynomials not in Π and free of x_i, $i > k$.

Let $S = S_1S_2J$. In

$$SF = S_1S_2MA_k^{h_k} + S_2S_1R$$

we substitute the expressions for S_2M and S_1R just described. There results an expression for SF as a linear combination of power products of $A_1, \ldots, A_k$. Those power products obtained from S_2S_1R are of degree less than h_k in A_k and therefore distinct from the power products obtained from

$$S_1S_2MA_k^{h_k}.$$

One can verify immediately that the expression for SF has the properties prescribed in the strengthened form of the conclusion to Lemma I.

4. A second lemma.

LEMMA II. *Let $\mathfrak{K}$ be a difference field of characteristic* 0, *A an algebraically irreducible difference polynomial of order and effective order n in the difference ring $\mathfrak{K}\{y\}$, and $C^{(0)}$ $(= A)$, $C^{(1)}, C^{(2)}, \ldots$, a characteristic sequence of a non-singular component Σ of $\{A\}$. There exist difference polynomials $M^{(0)}, M^{(1)}, M^{(2)}, \ldots$, of orders not exceeding $n, n+1, n+2, \ldots$, respectively, which are not in Σ, and for which each product $C^{(k)}, M^{(k)}$ $k = 0, 1, \ldots,$ is a linear combination of $A, A_1, \ldots, A_k$ with coefficients of order not exceeding $n+k$, while the coefficient of A_k is not in Σ.*

Proof. We choose $M^{(0)} = 1$. We suppose $M^{(1)}, \ldots, M^{(k-1)}$ to have been found and demonstrate the existence of $M^{(k)}$.

Since A_k has remainder 0 with respect to the chain $C^{(0)}, \ldots, C^{(k)}$ there is a relation

$$TC^{(k)} = JA_k + L^{(0)}C^{(0)} + \ldots + L^{(k-1)}C^{(k-1)}, \qquad J \notin \Sigma.$$

Multiplying both sides of this equation by $M^{(0)} \ldots M^{(k-1)}$, replacing each $M^{(i)}C^{(i)}$, $i = 0, \ldots, k-1$, by the appropriate linear combination of $A, \ldots, A_i$, and putting $TM^{(0)} \ldots M^{(k-1)} = M^{(k)}$ there results

$$M^{(k)}C^{(k)} = N^{(0)}A + \ldots + N^{(k)}A_k,$$

with $N^{(k)} = JM^{(0)} \ldots M^{(k-1)} \notin \Sigma$. Since the formal partial derivative $\partial A_k/\partial y_{n+k}$ is not in Σ, it is immediately seen by differentiation of $M^{(k)}C^{(k)}$ that $M^{(k)}$, too, is not in Σ. This proves Lemma II.

5. The preparation process. Let Σ be a reflexive prime difference ideal of order n in the ring $\mathfrak{K}\{y\}$, where the difference field $\mathfrak{K}$ is inversive and of characteristic 0, and let $F \in \mathfrak{K}\{y\}$ be of order $n+k$, $k > 0$. The following theorem provides two expressions for F in terms of the first polynomial A of the characteristic set of Σ.

PREPARATION THEOREM. *There exist difference polynomials S, T, of order at most $n + k$, which are not in Σ, and positive integers s, t such that*

$$SF = \sum_{i=1}^{s} L^{(i)} A^{p_{i,0}} A_1^{p_{i,1}} \dots A_k^{p_{i,k}}, \tag{3}$$

$$TF = \sum_{i=1}^{t} N^{(i)} A^{q_{i,0}} A_1^{q_{i,1}} \dots A_k^{q_{i,k}}. \tag{4}$$

Here the $p_{i,j}$ are non-negative integers, no two sets $p_{a,j}$, $p_{b,j}$ $(a \neq b, j = 0, \dots, k)$ are identical, the $q_{i,j}$ have a similar description, and the $L^{(i)}$ are difference polynomials of order not exceeding $n + k$. Those $L^{(i)}$ which are coefficients of terms whose power products in $A, A_1, \dots, A_k$ are of least weight for any positive value of the weight parameter not exceeding 1 are not in Σ. Also the $N^{(i)}$ are difference polynomials of order not exceeding $n + k$, while those $N^{(i)}$ which are coefficients of terms whose power products in $A, A_1, \dots, A_k$ are of least weight for any value of the weight parameter not less than 1 are not in Σ.

6. Proof of (3). Let $C^{(0)}$ $(= A)$, $C^{(1)}, \dots, C^{(k)}$ be the first $k + 1$ polynomials of a characteristic sequence of Σ. They are the characteristic set of the prime ideal

$$\Sigma' = \Sigma \cap \mathfrak{K}[y, y_1, \dots, y_{n+k}].$$

According to Lemma I there exists a relation

$$RF = \sum_{i=1}^{r} P^{(i)} (C^{(0)})^{a_{i,0}} \dots (C^{(k)})^{a_{i,k}}, \tag{5}$$

where the $a_{i,j}$ have a description similar to that of the $p_{i,j}$ of (3), and R and the $P^{(i)}$ are polynomials of $\mathfrak{K}[y, y_1, \dots, y_{n+k}]$ which are not in Σ', hence not in Σ.

Let polynomials $M^{(i)}$ be chosen in accordance with Lemma II. Putting $a = \max(a_{i,j})$, let $Q = R(M^{(0)} \dots M^{(k)})^a$. Then, using (5) and substituting for the $C^{(i)}M^{(i)}$ the linear combinations described in Lemma II, one finds a relation

$$QF = \sum_{i=1}^{q} J^{(i)} A^{r_{i,0}} A_1^{r_{i,1}} \dots A_k^{r_{i,k}}, \qquad Q \notin \Sigma, \tag{3*}$$

where the $r_{i,j}$ have a description similar to that of the $p_{i,j}$ of (3). We shall show that those $J^{(i)}$ which are coefficients of terms whose power products in the A_i are of least weight for any positive value of the weight parameter less than 1 are not in Σ.

To the ith term of (5) we assign a weight function $w_i(\theta) = a_{i,0} + a_{i,1}\theta + \dots + a_{i,k}\theta^k$. We consider a positive value $\tau < 1$ of the weight parameter. Then, upon the substitutions prescribed above to obtain (3*), the ith term of (5) gives rise to a term $T^{(i)}$ of the form

$$K^{(i)} A_0^{a_{i,0}} \dots A_k^{a_{i,k}}.$$

$K^{(i)} \notin \Sigma$, and other terms which, because their power products are formed from that of $T^{(i)}$ by replacing one or more of the A_j by lower transforms of A, are of greater weight than $T^{(i)}$. The coefficients of these terms may be in Σ.

If, in particular, the ith term of (5) is one of the terms whose power products are of least weight for the value τ of the weight parameter, then no term of (5) will yield a term of lower weight than $T^{(i)}$. Those terms of (5) of the same weight as $T^{(i)}$ will yield terms of this weight but with different power products. Hence, $T^{(i)}$ will actually be a term of (3*), and one of least weight for the value τ of the weight parameter. Clearly, all terms of (3*) of least weight for the value τ of the weight parameter arise in the same way as $T^{(i)}$ and, hence, have coefficients which are not in Σ. Hence, (3*) has the property claimed for (3) in the preparation theorem, except possibly for terms of least weight for the value 1 of the weight parameter, that is, for terms of least degree in the A_i.

Because the weight function is continuous, at least one of the terms of least degree of (3*) is a term of least weight for a value of the weight parameter less than 1, and hence, has a coefficient which is not in Σ.

Suppose the term

$$J^{(i)}A^{r_{i,0}}A_1^{r_{i,1}}A_k^{r_{i,k}}$$

is a term of (3*) which is of least degree, but that $J^{(i)} \in \Sigma$. Following the procedure used to obtain (3*) we find an equation

$$PJ^{(i)} = \sum_{j=1}^{p} H^{(j)}A^{s_{j,0}}A_1^{s_{j,1}}\dots A_k^{s_{j,k}}, \qquad P \notin \Sigma. \tag{6}$$

No term of the right-hand side of (6) is free of the A_i. For such a term would be of least weight for values of the weight parameter less than 1; hence, its coefficient would not be in Σ. This would yield a contradiction to the fact that $J^{(i)} \in \Sigma$.

Let $Q' = QP$. Multiplying both sides of (3*) by P, and substituting for $PJ^{(i)}$ from (6) we obtain an expression for $Q'F$ whose terms are those of the right-hand side of (3*) multiplied by P, except that the ith term of (3*) has been replaced by terms whose power products are multiples of its power product by power products of positive degree. Consequently, the terms of this expression which are of least weight for values of the weight parameter less than 1 have coefficients which are not in Σ, and the number of terms of least degree with coefficients in Σ is less than the number of such terms in (3*). Continuing the procedure just described, we obtain the equation (3).

7. Proof of (4). We define the difference field $\mathfrak{K}'$ to be the difference field whose elements are those of $\mathfrak{K}$ with the same addition and multiplication operations, but with transforming defined to be the inverse of the transforming operation of $\mathfrak{K}$. Let $\mathfrak{L}$ denote the inversive extension of $\mathfrak{K}\langle\alpha\rangle$,

where α denotes a generic zero of Σ. We define $\mathfrak{L}'$ to have the same relation to $\mathfrak{L}$ as $\mathfrak{K}'$ to $\mathfrak{K}$. Then $\mathfrak{L}'$ is an extension of $\mathfrak{K}'$.

Let $P \in \mathfrak{K}\{y\}$ be of order at most $n + k$. We define P' to be the polynomial of $\mathfrak{K}'\{z\}$ obtained by replacing each y_i in P by z_{n+k-i}. The operation $'$ produces a one-one correspondence between difference polynomials of order at most $n + k$ in $\mathfrak{K}\{y\}$ and in $\mathfrak{K}'\{z\}$. In particular, $B = (A_k)'$ is of order n, and B_h, $0 < h < k$, is $(A_{k-h})'$. α_{n+k} (where the subscript refers to transforming in $\mathfrak{L}$) is a generic zero of a reflexive prime difference ideal Σ' of $\mathfrak{K}'\{z\}$ whose characteristic set begins with B. The correspondence produced by $'$ maps the polynomials of Σ of order not exceeding $n + k$ onto the polynomials of Σ' of order not exceeding $n + k$.

Since (3) has been established, we know that there exists a relation

$$T'F' = \sum_{i=1}^{t} N^{(i)\prime} B^{e_{i,k}} B_1^{e_{i,k-1}} \dots B_k^{e_{i,0}}, \tag{4'}$$

meeting requirements corresponding to those imposed on (3). Now (4′) yields (4) on application of the inverse of the correspondence produced by $'$. Clearly, $T \notin \Sigma$. It remains only to show that the $N^{(i)}$ have the stated property.

Let $T^{(i)}$ denote the ith term of (4) and

$$w_i(\theta) = q_{i,0} + q_{i,1}\theta + \dots + q_{i,k}\theta^k$$

its weight function. The ith term $T^{(i)\prime}$ of (4′) has the weight function

$$v_i(\theta) = q_{i,0}\theta^k + \dots + q_{i,k} = \theta^k w_i(1/\theta).$$

Hence, if $T^{(i)}$ is a term of (4) of least weight for the value $\tau > 1$ of the weight parameter, then $T^{(i)\prime}$ is a term of least weight of (4′) for the value $1/\tau < 1$ of the weight parameter. Then $N^{(i)\prime} \notin \Sigma'$, so $N^{(i)} \notin \Sigma$. This completes the proof of the preparation theorem.

8. Proof of Equivalence. We now assume that F and A as described in § 5 also satisfy the conditions (a), (b), and (c) stated in the main theorem. As in that theorem, F is to be irreducible and vanish on the irreducible manifold $\mathfrak{M}$ of order n. As in the discussion of the preparation theorem, $\mathfrak{K}$ is assumed to be inversive, and A is chosen to be the first polynomial of the characteristic set of the reflexive prime difference ideal Σ with manifold $\mathfrak{M}$. In addition, we assume that F is of order and effective order $n + k$ with $k > 0$. It will be shown that there exists a power product U of the A_i which is of positive degree and is free of A and A_k such that

(α) *the right-hand side of* (3) *contains a term* YU, $Y \notin \Sigma$, *which is the term of least weight in the* A_i *for each positive value of the weight parameter not greater than* 1,

(β) *the right-hand side of* (4) *contains a term* ZU, $Z \notin \Sigma$, *which is the term of least weight in the* A_i *for each value of the weight parameter not less than* 1.

We follow the notation used in the statement of the main theorem. Because $\partial A/\partial y$ and $\partial A/\partial y_n$ are not in Σ, A^* is a polynomial of first degree which effectively contains z and z_n. S^* and T^* are in $\mathfrak{K} <\alpha>$. We find

$$F^* = \sum{}' \; (L^{i*}/S^*)(A^*)^{p_{i,0}} \ldots (A_k{}^*)^{p_{i,k}}, \tag{7}$$

$$F^* = {}_1\sum{}'' \; (N^{i*}/T^*)(A^*)^{q_{i,0}} \ldots (A_k{}^*)^{q_{i,k}}, \tag{8}$$

where Σ' and Σ'' are taken over those terms of (3) and (4) respectively which are of least degree in the A_i. It follows from the preparation theorem that the L^{*i} and N^{*i} appearing in these sums are elements of $\mathfrak{K} <\alpha>$.

Suppose that Σ', say, consists of more than one term. Then the homogeneous polynomial

$$\sum{}' \; (L^{*i}/S^*)u^{p_{i,0}} \ldots u_k^{p_{i,k}}$$

of the difference ring $\mathfrak{K} <\alpha>\{u\}$ is not a product of irreducible factors of effective order 0, and hence has a solution $u = \beta$, $\beta \neq 0$. The polynomial $A^* - \beta$ has a solution $z = \gamma$. Then γ is a solution of F^* but not of A^*, contrary to condition (c) of the main theorem. Hence, Σ' consists of just one term. Similarly, Σ'' consists of just one term. Because the left-hand sides of (7) and (8) are identical, and the coefficients on the right-hand sides are in $\mathfrak{K} <\alpha>$, it is clear that the same power product of the A^*_i occurs in each of these terms. We denote by U the corresponding power product of the A_i.

Consider a value $\tau < 1$ of the weight parameter. Let w_i be the weight of the term $T^{(i)} = L^{(i)}A^{p_{i,0}} \ldots A_k{}^{p_{i,k}}$ of (3). Upon substituting $y = z + \alpha$ the power product

$$A^{p_{i,0}} \ldots A_k^{p_{i,k}}$$

yields a term with power product

$$z_n^{p_{i,0}} \ldots z_{n+k}^{p_{i,k}}$$

of weight $\tau^n w_i$ and other terms whose weights are greater, since their power products are formed by replacing transforms of z in the indicated term by lower transforms.

If, in particular, $T^{(i)}$ is a term of (3) of least weight, $L^{(i)}$ is not in Σ, so that $T^{(i)}$ itself will yield a term with the above power product. Terms of greater weight than $T^{(i)}$ must yield only power products of z and its transforms of weight greater than $\tau^n w_i$. If

$$T^{(j)} = L^{(j)}A^{p_{j,0}} \ldots A_k^{p_{j,k}}, \qquad j \neq i,$$

is also of weight w_i it yields a distinct power product of weight $\tau^n w_i$ and other power products of greater weight. Hence, one power product of weight $\tau^n w_i$ appears in the polynomial $\bar{F}$ of the main theorem for each term of (3) of least weight, and these power products are of least weight in $\bar{F}$. Since, by hypothesis, there is a unique term of $\bar{F}$ of least weight, (3) contains a unique term of least weight for the value τ of the weight parameter. By continuity of the weight function this term is the same for all values of the weight

parameter less than 1. It follows at once that it is the term of (3) of least degree.

Let the weight parameter be $\tau > 1$. Let the term

$$N^{(i)}A^{e_{i,0}} \ldots A_k^{e_{i,k}}$$

of (4) have weight w_i. Upon substituting $y = z + \alpha$ into the power product

$$A^{e_{i,0}} \ldots A_k^{e_{i,k}}$$

there results a term with power product

$$z^{e_{i,0}} \ldots z_k^{e_{i,k}}$$

of weight w_i and other terms whose weights are greater since their power products are formed by replacing transforms of z in the indicated term by higher transforms. It follows, as above, that the term of (4) of least degree is the unique term of least weight for all values of the weight parameter exceeding 1.

Not every term on the right-hand side of (3) actually contains A, since F has no factors of order n, and A does not divide S. For sufficiently small values of the weight parameter a term free of A is certainly of lower weight than a term which contains A. Hence, U is free of A. Using (4) and considering large values of the weight parameter, we find that U is free of A_k. This completes the proof of the statements made at the beginning of this section.

It is easy, but unnecessary for the proof of the main theorem, to show that (α) and (β) are equivalent to (a), (b), and (c). Let (α) and (β) hold. For positive values less than 1 of the weight parameter, $\bar{U}$ (for explanation of the notation see the paragraph preceding the statement of the main theorem) contains a unique term of least weight. It follows from (3) and (α) that this term furnishes the unique term of least weight in $\bar{F}$. In a similar way it follows from (4) and (β) that $\bar{F}$ contains a unique term of least weight for values of the weight parameter exceeding 1. Hence, (a) and (b) hold. From either (3) or (4) there results $F^* = \gamma U^*$, $\gamma \in \mathfrak{K}\langle\alpha\rangle$. Since, clearly, U^* is a product of powers of transforms of A^*, this implies (c).

9. Completion of the proof. If $\mathfrak{M}$ is not an essential singular manifold of F there exist, as we shall see, certain formal power series solutions of F and its transforms. But we shall also see that such solutions cannot exist when $k = 2$ and the conditions (α) and (β) hold. These facts establish the main theorem. Throughout this work we maintain the restrictions of § 8. These restrictions are removed in § 14.

10. Existence of series solutions. We suppose that $\mathfrak{M}$ is not an essential singular manifold of F. Then there exists a reflexive prime difference ideal Λ containing F and properly contained in Σ. According to Lemma IV of **(3)**, Λ is of effective order greater than n, so that $A \notin \Lambda$. For any integer $r > k$

we define Σ_r and Λ_r to be the intersections of Σ and Λ respectively with the ring $\mathfrak{K}_r = \mathfrak{K}\,[y, \ldots, y_{n+r}]$. Let $\mathfrak{G} = \mathfrak{K}<\alpha>$. Λ_r generates an ideal in $\mathfrak{G}\,[y, \ldots, y_{n+r}]$ whose radical is the intersection of prime components at least one of which admits the solution $y_i = \alpha_i$ $(i = 0, 1, \ldots, n+r)$. Let Λ_r' be such a component. Since Λ_r' and Λ_r have the same dimension (**5**, vol. 2, p. 69), $\Lambda_r' \cap \mathfrak{K}_r = \Lambda_r$. Hence, $A \ldots A_r \notin \Lambda_r'$. Then (**4**, p. 526) Λ_r' admits a solution not annulling $A \ldots A_r$

$$y_i = \alpha_i + g_i(h), \qquad i = 0, 1, \ldots, n+r, \tag{9}$$

where h is transcendental over $\mathfrak{G}$ and the $g_i(h)$ are formal series in positive integral powers of h with coefficients algebraic over $\mathfrak{G}$.

11. Non-existence of series solutions. We now suppose that $k = 2$, and that the conditions (α) and (β) hold. We assume the existence of the solutions (9) and obtain a contradiction if r is sufficiently large.

If the series (9) is substituted into a polynomial P of $\mathfrak{K}_r$ there results a series in non-negative integral powers of h. The term of zero degree in this series results from the substitution of the α_i into P and, hence, is 0 if and only if $P \in \Sigma$. In particular, the series obtained from $A, A_1, \ldots, A_r$ are not 0 but begin with terms of positive degree. We denote the series obtained from A_i by

$$k_i(h) = a_i h^{s_i} + \ldots, \qquad i = 0, \ldots, r, \tag{10}$$

where the a_i are algebraic over $\mathfrak{G}$ and not 0, and the s_i are positive. Substitution of (9) into $F, \ldots, F_{r-k}$ $(= F_{r-2})$ gives 0, while substitution into S, T, or the coefficient of the term of least degree on the right-hand side of (3) or of (4) results in a series whose term of zero degree is not 0.

12. We consider the power product U of the A_i described in § 8. Since $k = 2$, $U = A_1^d$, $d > 0$. The numbering of the terms of the right-hand sides of (3) and (4) is to be chosen so that those terms whose power products are of degree less than d in A_1 precede the remaining terms, and the term with power product U is last. Let s' and t' denote the number of terms on the right-hand sides of (3) and (4) respectively whose power products are of degree less than d in A_1. Since not every term on the right-hand side of (3) or of (4) has the factor A_1, $1 \leqslant s' < s$, $1 \leqslant t' < t$.

Let

$$\begin{aligned} p_i(\theta) &= p_{i,0} + (p_{i,1} - d)\theta + p_{i,2}\theta^2, \qquad 1 \leqslant i \leqslant s', \\ q_i(\theta) &= q_{i,0} + (q_{i,1} - d)\theta + q_{i,2}\theta^2, \qquad 1 \leqslant i \leqslant t'. \end{aligned} \tag{11}$$

Since the power product A_1^d is of lower weight than the other power products on the right-hand side of (3) for positive values of the weight parameter not exceeding 1, $p_i(\theta) > 0$, $0 < \theta \leqslant 1$. Since $p_{i,1} - d < 0$, $1 \leqslant i \leqslant s'$, it follows that $p_{i,0} > 0$ for these i. Then each $p_i(\theta)$ is bounded away from 0 on the interval $0 \leqslant \theta \leqslant 1$. Similarly, $q_i(\theta) > 0$, $\theta \geqslant 1$, whence it follows that the

$q_i(\theta)$ are bounded away from 0 on this interval, and that the $q_{i,2}$, $1 \leqslant i \leqslant t'$, are all positive. Let $m > 0$ be such that

$$(12) \qquad \begin{aligned} p_i(\theta) &> m, & 0 \leqslant \theta \leqslant 1, & \quad 1 \leqslant i \leqslant s', \\ q_i(\theta) &> m, & \theta \geqslant 1, & \quad 1 \leqslant i \leqslant t'. \end{aligned}$$

We define a to be the maximum of the quotients $(d - q_{i,1})/q_{i,2}$, $1 \leqslant i \leqslant t'$, and b to be the maximum of the $p_{i,2}$, $1 \leqslant i \leqslant s'$, $q_{i,2}$, $1 \leqslant i \leqslant t'$. Then $a, b > 0$. Let $c = m/ab$, $d = a/c$. We shall obtain a contradiction from (9) with $r \geqslant d + 2$.

13. Upon substituting the series (9) into the right-hand side of (3) the result must be zero. This can only be so if the power product of some term of (3) other than the last term produces an expansion beginning with a power of h not higher than that with which the expansion of the last term begins. Since the last term is YA_1^d, $Y \notin \Sigma$, this means that there is an integer j_0 such that

$$p_{j_0,0}s_0 + (p_{j_0,1} - d)s_1 + p_{j_0,2}s_2 \leqslant 0.$$

From the definition of s' it is clear that $1 \leqslant j_0 \leqslant s'$. By applying similar reasoning to the right-hand side of (4) and to the transforms of orders not exceeding $r - 2$ of the right-hand sides of (3) and of (4) it follows that there exist integers $j_i, k_i, 0 \leqslant i \leqslant r - 2$, such that

$$(13) \qquad \begin{aligned} &1 \leqslant j_i \leqslant s', \qquad 1 \leqslant k_i \leqslant t'; \\ &p_{j_i,0}s_i + (p_{j_i,1} - d)s_{i+1} + p_{j_i,2}s_{i+2} \leqslant 0; \\ &q_{k_i,0}s_i + (q_{k_i,1} - d)s_{i+1} + q_{k_i,2}s_{i+2} \leqslant 0. \end{aligned}$$

Let $t_i = s_{i+1}/s_i > 0$, $(i = 0, \ldots, r - 1)$. Then (13) yields, for $0 \leqslant i \leqslant r - 2$,

$$(14) \qquad \begin{aligned} &p_{j_i,0} + (p_{j_i,1} - d)t_i + p_{j_i,2}t_it_{i+1} \leqslant 0; \\ &q_{k_i,0} + (q_{k_i,1} - d)t_i + q_{k_i,2}t_it_{i+1} \leqslant 0. \end{aligned}$$

It follows from (12) that for each i, $0 \leqslant i \leqslant r - 2$, either $0 < t_i \leqslant 1$, and

$$p_{j_i,0} + (p_{j_i,1} - d)t_i + p_{j_i,2}t_i^2 > m;$$

or $t_i > 1$, and

$$q_{k_i,0} + (q_{k_i,1} - d)t_i + q_{k_i,2}t_i^2 > m.$$

From whichever of these inequalities is applicable it follows by subtraction of the corresponding inequality of (14) that either

$$p_{j_i,2}t_i(t_i - t_{i+1}) > m,$$

or

$$q_{k_i,2}t_i(t_i - t_{i+1}) > m.$$

In either case

$$(15) \qquad t_i(t_i - t_{i+1}) > m/b, \qquad 0 \leqslant i \leqslant r - 2.$$

Now (14) yields

$$(q_{k_i,1} - d) + q_{k_i,2}t_{i+1} \leqslant 0, \qquad 0 \leqslant i \leqslant r-2$$

since, for every i,

$$q_{k_i,0} \geqslant 0.$$

Thus,

$$t_{i+1} \leqslant (d - q_{k_i,1})/q_{k_i,2} \leqslant a, \qquad 0 \leqslant i \leqslant r-2. \tag{16}$$

From (15) and (16) there results $t_1 \leqslant a$, $t_i - t_{i+1} > m/ba = c$, $1 \leqslant i \leqslant r-2$. Hence, $t_{r-1} < a - (r-2)c \leqslant a - dc = 0$. This is the desired contradiction.

14. Removal of the restrictions. It remains only to prove the main theorem without the restrictions of § 9. Let $\bar{\mathfrak{K}}$ be the inversive extension of $\mathfrak{K}$, C the polynomial of order n in $\bar{\mathfrak{K}}\{y\}$ and G the polynomial of order $n+k$ in $\bar{\mathfrak{K}}\{y\}$ whose transforms of the appropriate orders are A and F respectively. Let $\mathfrak{M}'$ be the irreducible manifold over $\bar{\mathfrak{K}}\{y\}$ with generic zero α, and Σ' the reflexive prime difference ideal of $\bar{\mathfrak{K}}\{y\}$ with manifold $\mathfrak{M}'$. Then $\mathfrak{M}'$ is of order n, each irreducible factor of G is of order and effective order $n+k$, and $G \in \Sigma'$.

Let H be an irreducible factor of G which is in Σ'. Then, in the notation of the main theorem, the polynomial consisting of the terms of $\bar{H}$ of least weight for some value of the weight parameter is a factor of the polynomial consisting of the terms of $\bar{G}$ of least weight for this value of the weight parameter. The latter polynomial is an inverse transform of the polynomial consisting of the terms of $\bar{F}$ of least weight.

The first polynomial D of a characteristic set of Σ' divides C. Let $C = PD$. Since α is not a solution of $\partial C/\partial y_n$, $P \notin \Sigma'$. Then $D^* = \gamma C^*$, $\gamma \in \bar{\mathfrak{K}}\langle\alpha\rangle$, $\gamma \neq 0$; and C^* is an inverse transform of A^*.

The preceding statements show that $\mathfrak{M}'$ and H satisfy the conditions (a), (b), and (c), so that $\mathfrak{M}'$ is an essential singular manifold of H according to the restricted case of the main theorem. Hence, there is a polynomial $Q \in \bar{\mathfrak{K}}\{y\}$ such that $Q \notin \Sigma'$, and, if $E \in \Sigma'$, QE vanishes on the manifold of H.

To each irreducible factor of G there corresponds a polynomial with the properties of Q. For this has just been proved for factors which vanish on $\mathfrak{M}'$, and it is evident for other factors. Let R be the product of these polynomials. Then $R \notin \Sigma'$, and, if $E \in \Sigma'$, RE vanishes on the manifold of G. Some transform S of R is in $\mathfrak{K}\{y\}$. Since $\Sigma \leqslant \Sigma'$, $S \notin \Sigma$, and, if $E \in \Sigma$, SE vanishes on the manifold of F. This proves that $\mathfrak{M}$ is a component of $\{F\}$, and, indeed, since its effective order is less than that of F, an essential singular manifold of F. The proof of the main theorem is complete.

15. Constructive methods. It is possible to determine by actual construction whether or not conditions (a), (b), and (c) are satisfied, provided one knows the first k polynomials of a characteristic sequence of Σ. (For the

meaning to be given to "characteristic sequence" if Σ is not of equal order and effective order, see (**2**, footnote 7).) In fact, it was shown in (**2**, p. 447) that one can determine constructively whether or not (a) and (b) hold. But if (a) and (b) hold, (c) is true if and only if F^* is a product of powers of transforms (including, possibly, inverse transforms) of A^* and a factor in $\mathfrak{R}<\alpha>$. For, on the one hand, this condition clearly implies (c). On the other hand, if (a), (b), and (c) hold it follows from (α) and (β) under the conditions of the restricted form of the main theorem, and from this special case and the reasoning of § 14 in the general case, that F^* is such a product. There is no difficulty in determining constructively whether or not F^* is a product of this type. It follows, in particular, that, if $k = 2$, one can determine constructively, under the stated limitation, whether or not $\mathfrak{M}$ is an essential singular manifold of F.

References

1. J. F. Ritt, *Differential algebra*, Coll. publ. Amer. Math. Soc., *33*.
2. R. M. Cohn, *Singular manifolds of difference polynomials*, Ann. Math., *53* (1951), 445–63.
3. ——— *Extensions of difference fields*, Amer. J. Math., *74* (1952), 507–30.
4. ——— *Essential singular manifolds of difference polynomials*, Ann. Math., *57* (1953), 524–30.
5. W. V. D. Hodge and D. Pedoe, *Methods of algebraic geometry* (Cambridge University Press).

Rutgers University

SUBSPACES OF A GENERALIZED METRIC SPACE

H. A. ELIOPOULOS

Introduction. In a paper published in 1956, Rund **(4)** developed the differential geometry of a hypersurface of $n-1$ dimensions imbedded in a Finsler space of n dimensions, considered as locally Minkowskian.

The purpose of the present paper is to provide an extension of the results of **(4)** and thus develop a theory for the case of m-dimensional subspaces imbedded in a generalized (Finsler) metric space.

We consider an n-dimensional differentiable manifold X_n and we restrict our attention to a suitably chosen co-ordinate neighbourhood of X_n in which a co-ordinate system x^i $(i = 1, 2, \ldots, n)$, is defined. A system of equations of the type $x^i = x^i(t)$ defines a curve C of X_n, the tangent vector dx^i/dt of which is denoted by $\dot{x}^i$. We say that the manifold X_n is endowed with a locally Minkowskian (Finsler) metric, if the length of an arc of the curve C between two points P_1 and P_2 of C, corresponding to parameter values t_1 and t_2, is defined by an integral of the type

$$\int_{t_1}^{t_2} F(x^i, \dot{x}^i)dt,$$

where the function $F(x^i, \dot{x}^i)$ is continuous and continuously differentiable up to any required order in all its arguments, and also positively homogeneous of the first degree in the $\dot{x}^i$.

Defining the metric tensor of X_n by

$$g_{ij}(x, \dot{x}) = \frac{1}{2}\frac{\partial^2 F^2(x, \dot{x})}{\partial \dot{x}^i \partial \dot{x}^j}, \qquad g^{ik}(x, \dot{x})g_{ih}(x, \dot{x}) = \delta^k_h,$$

we can put

$$F^2(x, \dot{x}) = g_{ij}(x, \dot{x})\dot{x}^i\dot{x}^j;$$

F must satisfy a third condition,

$$g_{ij}(x, \dot{x})\xi^i\xi^j > 0,$$

for all $\dot{x}^i$ and all ξ^i, provided not all ξ^i are equal to zero.

From Euler's theorem on homogeneous functions we have

$$\frac{\partial g_{ij}(x, \dot{x})}{\partial \dot{x}^k}\dot{x}^i = 0, \qquad \frac{\partial^2 g_{ij}(x, \dot{x})}{\partial x^h \partial \dot{x}^k}\dot{x}^i = 0.$$

Received April 9, 1958. The present paper is based on a thesis submitted at the University of Toronto for the degree of Doctor of Philosophy. The author wishes to express his sincere appreciation to Professor H. Rund for direction and advice in the course of this investigation, and to Professor G. F. D. Duff for valuable comments. The National Research Council of Canada supported the research by a fellowship.

We also define the generalized Christoffel symbols of the first and second kind by the relations

$$\begin{Bmatrix} i \\ hk \end{Bmatrix}_{(x,\dot{x})} = g^{ij}(x, \dot{x})[hk, j]_{(x,\dot{x})},$$

$$[hk, j]_{(x,\dot{x})} = \tfrac{1}{2}\left(\frac{\partial g_{kj}(x, \dot{x})}{\partial x^h} + \frac{\partial g_{hj}(x, \dot{x})}{\partial x^k} - \frac{\partial g_{hk}(x, \dot{x})}{\partial x^j}\right).$$

Let C be a continuous and continuously differentiable curve. At each point P of C, with co-ordinates x^k, a Minkowskian tangent space $T_n(P)$ is defined by $F(x^k, \dot{x}^k)$. We consider an arbitrary vector field $X^i(x^k)$ along C such that in each $T_n(P)$ a vector X^i is defined. Let Q be a neighbouring point with co-ordinates $x^k + dx^k$ on C, such that the arc length $PQ = ds$. The covariant differential DX^i of X^i at P for the transition from P to Q is then defined by

$$DX^i = \left(\frac{\partial X^i}{\partial x^k} + P^i_{hk}(x, x')X^h\right)dx^k, \tag{A.3}$$

where

$$P^i_{hk}(x, x') = \begin{Bmatrix} i \\ hk \end{Bmatrix}_{(x,x')} - \tfrac{1}{2}\, g^{im}(x, x')\, \frac{\partial g_{hm}(x, x')}{\partial x'^l} \begin{Bmatrix} l \\ pk \end{Bmatrix}_{(x,x')} x'^p,$$

and $x'^i = dx/ds$.

We note that (A.3) depends only on the vector X^i and the displacement PQ for which it has been defined, and not on the curve C passing through P and Q. On the other hand, the covariant derivative of X^i with respect to x^k is given by

$$X^i_{;k} = \frac{\partial X^i}{\partial x^k} + P^{*i}_{hk}(x, x')x^h, \tag{A.4}$$

where **(5)**

$$P^*_{ij,k} = g_{kh}P^{*h}_{ij} = [ij, k] - \tfrac{1}{2}\left(\frac{\partial g_{kj}}{\partial x'^l} P^l_{ik} + \frac{\partial g_{ki}}{\partial x'^k} P^l_{jk} - \frac{\partial g_{ij}}{\partial x'^h} P^l_{hk}\right) x'^k.$$

Consider a continuous curve C of X_n, which lies in some two-dimensional subspace X_2 of X_n, and let the parameters of X_2 be u and v. The parametric curves u = const. and v = const. may cut C in an arbitrary manner. Two directions

$$\xi^k = \frac{\partial x^k}{\partial u}, \qquad \eta^k = \frac{\partial x^k}{\partial v}$$

are defined at each point of C, and they represent the directions of the tangents to the co-ordinate curves. Then, for a vector field $X^i(x^k)$, we have in the X_2,

$$\frac{DX^i}{Du} = X^i_{;k}\xi^k, \qquad \frac{DX^i}{Dv} = X^i_{;k}\eta^k,$$

and thus, we obtain the commutation formula **(6)**,

$$\frac{D^2X^i}{DvDu} - \frac{D^2X^i}{DuDv} = (X^i_{;nm} - X^i_{;mn})\xi^n\xi^m + X^i_{;n}(\xi^n_{;m}\eta^m - \eta^n_{;m}\xi^m). \tag{A.6}$$

If we use the relation

$$\frac{\partial \xi^k}{\partial v} = \frac{\partial \eta^k}{\partial u},$$

we reduce (A.6) to

$$\frac{D^2X^i}{DvDu} - \frac{D^2X^i}{DuDv} = (X^i_{;nm} - X^i_{;mn})\xi^n\eta^m.$$

Introducing the expression

$$K^i_{.hmn}(x, x') = \frac{\partial P^{*i}_{hn}}{\partial x^m} - \frac{\partial P^{*i}_{hm}}{\partial x^n} + P^{*i}_{sm}P^{*s}_{hn} - P^{*i}_{sn}P^{*s}_{hm} + \left(\frac{\partial P^{*i}_{hn}}{\partial x'^l}\frac{\partial x'^l}{\partial x^m} - \frac{\partial P^{*i}_{hm}}{\partial x'^l}\frac{\partial x'^l}{\partial x^n}\right)_{(x,x')}, \tag{A.7}$$

which we call the relative curvature tensor in view of the derivative $\partial x'^l/\partial x^m$ which appears in it, we may obtain the commutation relation

$$X^i_{;nm} - X^i_{;mn} = K^i_{.hmn}X^h.$$

We also define a covariant curvature tensor from the relation

$$K_{ihmn}(x, x') = g_{ij}(x, x')\, K^j_{.hmn}(x, x');$$

then, if $Y_i(x^k)$ are the covariant components of the vector field, we may obtain the relation

$$Y_{i,mn} - Y_{i,nm} = -K^h_{.imn}Y_h. \tag{A.8}$$

1. Generalities. Consider a differentiable subspace of m dimensions F_m, imbedded in a locally Minkowskian (Finsler) space F_n, where $m < n$. Let

$$x^i = x^i(u^\alpha), \qquad (i = 1 \ldots n, \alpha = 1 \ldots m), \tag{1.1}$$

be the equations defining F_m. We assume that the Jacobian matrix

$$(X^i_\alpha) = \left(\frac{\partial x^i}{\partial u^\alpha}\right)$$

is of rank m.

If the co-ordinate curves are regarded as curves of the F_m, then their tangents are given by

$$X^i_\alpha = \frac{\partial x^i}{\partial u^\alpha}$$

and at each point P of F_m we have m independent vectors $\partial x^i/\partial u^\alpha$, which will span an m-dimensional plane $T_m(P) \subset T_n(P)$, where by $T_m(P)$ we mean the m-dimensional linear space tangent to F_m at P.

A vector X^i lies in F_m if $X^i \in T_m(P)$, which implies that it is of the form

$$X^i = U^\alpha \frac{\partial x^i}{\partial u^\alpha}. \tag{1.2}$$

F_m will be endowed with an induced metric

$$ds^2 = g_{\alpha\beta}(u, u')du^\alpha du^\beta$$

with fundamental tensor given by

$$g_{\alpha\beta}(u, u') = g_{ij}(x, x') \frac{\partial x^i}{\partial u^\alpha} \frac{\partial x^j}{\partial u^\beta}, \tag{1.3}$$

where the tangent u'^α to F_m, satisfies the relation

$$x'^i = X^i_\alpha u'^\alpha. \tag{1.4}$$

In general, we have to consider two sets of normals to F_m at a given point P of F_m. The first set is defined by the solutions n^i of the equations

$$n_i X^i_\alpha \equiv g_{ij}(x, n)n^j X^i_\alpha = 0. \tag{1.5}$$

These solutions are normalized by means of the relation

$$F(x, n) = 1 \quad \text{or} \quad g_{ij}(x, n)n^i n^j = 1. \tag{1.6}$$

Since the matrix $(X_\alpha{}^i)$ is of rank m, we have $n - m$ independent solutions and, therefore, $n - m$ independent normal vectors. They span a vector space at P, and any vector of this space will be a linear combination of the independent vectors spanning the space.

We may define a different set of normals in the following way. Let x'^i be an arbitrary but fixed direction tangential to F_m at P. A second set of normals can be defined by the solutions $n^*(x, x')$ of the equations

$$g_{ij}(x, x')n^{*j}(x, x')X^i_\alpha = 0. \tag{1.7}$$

The matrix $(X_\alpha{}^i)$ being of rank m, the system (1.7) admits $n - m$ independent solutions of the direction considered. We may write

$$\overset{*}{n}{}^i_{(\mu)} = \overset{*}{n}{}^i_{(\mu)}(x, x'), \qquad (\mu = 1 \dots n - m).$$

To each direction x' tangent to F_m at P corresponds a set of vectors $n^{*i}{}_{(\mu)}(x, x')$, and the totality of these sets, for the different x' at P, defines $n - m$ cones which are the normal cones of the subspace F_m at a given point. We must emphasize that the generators of the normal cones do not necessarily lie in the space spanned by the normals n at the same point. The concept of the normal cones for subspaces is an extension of the idea of a normal cone of a hypersurface F_{n-1} (4).

We assume as in the case of the $n(x)$, that $n^*(x, x')$ are normalized according to the relation

$$F(x, n^*(x, x')) = g_{ij}(x, n^*(x, x'))n^{*i}(x, x)n^{*j}(x, x') = 1. \tag{1.8}$$

We may also define $n - m$ tensors, independent of direction,

$$\gamma_{(\mu)\alpha\beta}(u) = g_{ij}(x, n_{(\mu)})X^i_\alpha X^j_\beta,$$

for the $n - m$ normals at P. Then we define the following sets of inverse projection parameters corresponding to $X_\beta{}^j$:

(1.9)
$$X^{\alpha}_{i}(x, x') = g_{ij}(x, x')g^{\alpha\beta}(u, u')X^{j}_{\beta},$$
$$Y^{\alpha}_{(\mu)i}(x) = g_{ij}(x, n_{(\mu)})\gamma^{\alpha\beta}_{(\mu)}(u)X^{j}_{\beta},$$

so that, in view of the equations (1.5), (1.7), we have

(1.10)
$$n^{*}_{(\mu)}X^{\alpha}_{i} = 0, \qquad Y^{\alpha}_{(\mu)i}n^{i}_{(\mu)} = 0,$$

and also

(1.10a)
$$X^{\alpha}_{i}X^{i}_{\beta} = \delta^{\alpha}_{\beta}, \qquad Y^{\alpha}_{(\mu)i}X^{i}_{\beta} = \delta^{\alpha}_{\beta}.$$

It is always possible to choose a set of $n - m$ orthogonal independent vectors $n^{*}(x, x')$. Indeed, for any vector of the space spanned by the $n^{*i}{}_{(\mu)}$ we have

$$N^{*i}(x, x') = \sum_{(\mu)} \lambda_{(\mu)}n^{*i}_{(\mu)}(x, x), \qquad (\mu = 1 \ldots n - m).$$

Let us consider a set of $n - m$ such vectors; we can write down the $n - m$ relations

$$N^{*i}_{(\nu)}(x, x') = \sum_{(\mu)} \lambda_{(\nu)(\mu)}n^{*i}_{(\mu)}(x, x'), \qquad (\nu, \mu = 1 \ldots n - m).$$

In order that $N^{*i}{}_{(\nu)}$ should be orthogonal (with respect to $g_{ij}\ (x, x')$) the functions $\lambda_{(\nu)(\mu)}$ must satisfy the relations

(1.11)
$$g_{ij}(x, x')N^{*i}_{(\nu)}N^{*j}_{(\sigma)} = \sum_{(\mu)} \sum_{(\kappa)} g_{ij}(x, x')\lambda_{(\nu)(\mu)}\lambda_{(\sigma)(\kappa)}n^{*i}_{(\mu)}(x, x')n^{*j}_{(\kappa)}(x, x') = \delta(\nu)(\sigma).$$

If we put

(1.12)
$$T_{(\mu)(\kappa)}(x, x') = g_{ij}(x, x')n^{*i}_{(\mu)}n^{*j}_{(\kappa)}, \qquad (\mu, \kappa = 1 \ldots n - m),$$

the equation (1.11) can be written

(1.13)
$$\sum_{(\mu)} \sum_{(\kappa)} T_{(\mu)(\kappa)}\lambda_{(\nu)(\mu)}\lambda_{(\sigma)(\kappa)} = 0, \qquad \text{for } \nu \neq \sigma.$$

Our problem reduces to finding $n - m$ sets of functions $\lambda_{(\nu)(\mu)}$ satisfying the equations (1.13).

It is known that, if in a projective $(n - 1)$-dimensional space we introduce homogeneous co-ordinates, the equation of a hyperquadric has the form

(1.13a)
$$a_{kl}x_{k}x_{l} = 0,$$

and the co-ordinates x_{k}, y_{l} of two points harmonically conjugate with respect to (1.13a) satisfy the relation

$$a_{kl}x_{k}y_{l} = 0,$$

(see **(1)** for the 2-dimensional case). The problem of finding sets of functions $\lambda_{(\mu)(\nu)}$ satisfying (1.13), is equivalent to the problem of finding the vertices of polyhedra self-polar with respect to

$$\sum_{(\mu)(\kappa)} T_{(\mu)(\kappa)}\lambda_{(\mu)}\lambda_{(\kappa)} = 0.$$

One vertex P_1 of such a polyhedron can be chosen arbitrarily in the space, but not on the quadric; a second vertex P_2, arbitrarily in the polar hyperplane

of P_1, but not on the quadric; a third vertex P_3, arbitrarily on the intersection of the polar planes of P_1, P_2, but not on the quadric, P_4 on the intersection of the polar planes of P_1, P_2, P_3, and so on. The last one will be on the intersection of the polar hyperplanes of all the previous points. Since $P_1, P_2, \ldots, P_{n-1}$ can be chosen with $n-m-1, n-m-2, \ldots, 1$ degrees of freedom respectively, there are

$$(n-m-1)+(n-m-2)+\ldots+1 = \tfrac{1}{2}(n-m)(n-m-1)$$

degrees of freedom in choosing the $n-m$ sets of functions λ.

The induced covariant derivative of the vector X^i can be defined just as for a hypersurface **(4)**. Let $x^i = x^i(s)$ be a curve C of F_m so that x'^i is tangent to F_m. We consider a continuous and continuously differentiable vector field tangent to F_m:

$$X^i(x^\epsilon) = X^i_\alpha U^\alpha(u^\beta). \tag{1.16}$$

The induced covariant derivative of the vector field along C in the space F_m, that is, the tensor defined by

$$U^\beta_{,\gamma}(u, u') = \frac{\partial U^\beta}{\partial u^\gamma} + P^{*\beta}_{\delta\gamma}(u, u')U^\delta \tag{1.17}$$

is given by projection onto F_m of the covariant derivative $X_{,k}{}^i$ of X^i with respect to F_n,

$$g_{ij}(x, x')X^j_\gamma X^k_\alpha X^i_{,k} = g_{\beta\gamma}(u, u')U^\beta_{,\alpha}, \tag{1.18}$$

where

$$X^i_{,k} = \frac{\partial X^i}{\partial x^k} + P^{*i}_{hk}(x, x')X^h. \tag{1.19}$$

One can prove easily that

$$g_{ij}(x, x')X^j_\gamma\left(\frac{\partial^2 x^i}{\partial u^\alpha \partial u^\beta} + P^{*i}_{hk}X^k_\alpha X^h_\beta\right) = P^*_{\beta\alpha,\gamma}(u, u'),$$

with

$$P^*_{\beta\alpha,\gamma}(u, u') = g_{\epsilon\gamma}P^\epsilon_{\beta\alpha}.$$

It is obvious that $P^{*\gamma}{}_{\beta\alpha}$ are symmetric in the lower indices, because $P^{*i}{}_{hk}$ are symmetric.

It is very easy to show that the quantities (1.17) *form the components of a tensor, in the sense indicated by their indices, under a transformation of the co-ordinates u^α of F_m*, Eliopoulos **(3)**.

Since the subspace F_m is endowed with a metric tensor $g_{\alpha\beta}(u, u')$, we can write immediately the Euler-Lagrange equations for the geodesics of that space

$$\frac{d^2u}{ds^2} + \begin{Bmatrix}\alpha \\ \beta\gamma\end{Bmatrix}_{(u,u')} \frac{du^\beta}{ds}\frac{du^\gamma}{ds} = 0,$$

where

$$\begin{Bmatrix} \alpha \\ \beta\gamma \end{Bmatrix}_{(u,u')}$$

are the intrinsic Christoffel symbols. We may also write

$$g_{\alpha\delta} \frac{du'^\alpha}{ds} + [\beta\gamma, \delta]_{(u,u')} \frac{du^\beta}{ds}\frac{du^\gamma}{ds} = 0,$$

or

$$g_{\alpha\delta} \frac{du'^\alpha}{ds} + P^{*\delta}_{\beta\delta} u'^\alpha u'^\beta = 0.$$

We immediately see that

$$\frac{\delta u'^\alpha}{\delta s} = 0$$

along a geodesic, that is, the geodesics are autoparallel curves.

2. Normal curvatures of F_m. We consider a curve C of F_m, $x_1 = x_1(s)$, passing through a given point P. We take the parameter s to be the arc-length, and the unit tangent vector to C at P will be denoted by x'^i. Let us assume, for the moment, that the vector field U^α of equations (1.16) coincides with the tangent vectors u'^α of C. If we denote covariant differentiations in F_m by δ, we obtain

$$\text{(2.1)} \qquad \frac{\delta u'^\alpha}{\delta s} = \frac{du'^\alpha}{ds} + P^{*\alpha}_{\beta\gamma}(u, u')u'^\beta u'^\gamma = \frac{du'^\alpha}{ds} + \begin{Bmatrix} \alpha \\ \beta\gamma \end{Bmatrix} u'^\beta u'^\gamma.$$

By using the expression of Dx'^i/Ds and differentiating $x' = X_\alpha{}^i u'^\alpha$ we find

$$\frac{Dx'^i}{Ds} = \frac{\partial^2 x^i}{\partial u^\alpha \partial u^\beta} u'^\alpha u'^\beta + X^i_\alpha \frac{\delta u'^\alpha}{\delta s} - X^i_\alpha P^{*\alpha}_{\beta\gamma} u'^\beta u'^\gamma + P^{*i}_{hk} x'^h x'^k.$$

If we put

$$\text{(2.3)} \qquad X^i_{\alpha\beta} = \frac{\partial^2 x^i}{\partial u^\alpha \partial u^\beta} - X^i_\gamma P^{*\gamma}_{\alpha\beta} + P^{*i}_{hk} X^h_\alpha X^k_\beta,$$

we may write

$$\text{(2.4)} \qquad \frac{Dx'^i}{Ds} = X^i_{\alpha\beta} u'^\alpha u'^\beta + X^i_\alpha \frac{\delta u'^\alpha}{\delta s}.$$

The expressions $X_{\alpha\beta}{}^i$ which are the components of a tensor, may be considered as the generalized covariant derivatives of the $X_\alpha{}^i$ with respect to u^β, in the sense used in **(4, 7)**. We note that $X_{\alpha\beta}{}^i$ are symmetric with respect to the lower indices.

The $X_{\alpha\beta}{}^i$ can be given the following geometric interpretation: We consider the geodesic $\bar{C}$, of the space F_n through the point P, tangent to the given direction x'^i. Let $\bar{x}^i = \bar{x}^i(s)$ be the equations of $\bar{C}$. We also consider a geodesic $\bar{C}$ of

the space F_n through the same point P, and tangent to x'^i. Let $\bar{x}^i = \bar{x}^i(s)$ be its equations. We choose two points one on $\bar{C}$ and the other on $\bar{C}$ corresponding to the same value of s, and in the neighbourhood of P. The coordinates of these points can be expended in Taylor series, for small values of s, so that

$$\bar{x}^i(s) = \bar{x}^i_P + \bar{x}'^i_P s + \tfrac{1}{2}\bar{x}''^i_P s^2 + \ldots$$
$$\bar{x}^i(s) = \bar{x}^i_P + \bar{x}'^i_P s + \tfrac{1}{2}\bar{x}''^i_P s^2 + \ldots,$$

where by $\bar{x}_P$, $\bar{x}_P$, etc., we mean the values of these functions at the point P. Then

$$\xi^i = \bar{x}^i - \bar{x}^i = \tfrac{1}{2}(\bar{x}''^i - \bar{x}''^i)s^2 + O(s^3)$$

because $\bar{x}_P{}^i = \bar{x}_P{}^i$ and $\bar{x}_P{}'^i = \bar{x}_P{}'^i$. From the equations of geodesics we have for $\bar{C}$

$$\frac{d\bar{x}'^i}{ds} = -\begin{Bmatrix} i \\ hk \end{Bmatrix}_{(a)} \bar{x}'^h\bar{x}'^k.$$

Also for $\bar{C}$, we have

$$\frac{D\bar{x}'^i}{Ds} = \frac{d\bar{x}'^i}{ds} + \begin{Bmatrix} i \\ hk \end{Bmatrix}_{(c)} \bar{x}'^h\bar{x}'^k,$$

and therefore

$$\xi^i = \tfrac{1}{2}\frac{D\bar{x}'^i}{Ds}s^2 + O(s^3).$$

In view of (2.4) applied to a geodesic of F_n we obtain

$$X^i_{\alpha\beta}(u, u')u'^\alpha u'^\beta = \lim_{s\to 0}\frac{2\xi^i}{s^2}.$$

We consider the formulae (1.5) and (1.7). Since $n_{(\mu)i}$ and $n^*{}_{(\nu)i}$ are solutions of the same linear equations, we may write

$$(2.8) \qquad n_{(\mu)i} = \sum_{(\nu)} p_{\mu\nu} \overset{*}{n}_{(\nu)i};$$

multiplying the above equations by $n_{(\mu)}{}^i$, and since $n_{(\mu)}{}^i$ are unit vectors, we obtain

$$\sum_{(\nu)} p_{\mu\nu}\overset{*}{n}_{(\nu)i} n^i_{(\mu)} = 1.$$

The equations (2.8) can also be written as

$$g_{ij}(x, n_{(\mu)})n^j_{(\mu)} = \sum_{(\nu)} p_{\mu\nu} g_{ij}(x, x')\overset{*}{n}{}^j_{(\nu)}$$

and if we multiply by $n^*{}_{(\lambda)}i$ we find

$$g_{ij}(x, n_{(\mu)})n^j_{(\mu)}\overset{*}{n}{}^i_{(\lambda)} = \sum_{(\nu)} p_{\mu\nu} g_{ij}(x, x')\overset{*}{n}{}^j_{(\nu)}\overset{*}{n}{}^i_{(\lambda)} = p_{\mu\lambda}\psi_\lambda,$$

since

$$g_{ij}(x, x')\overset{*}{n}{}^j_{(\nu)}\overset{*}{n}{}^i_{(\lambda)} = \delta^\nu_\lambda\psi_\lambda$$

(no summation over λ involved). The above relation may be written

$$\cos(n_{(\mu)}, \overset{*}{n}_{(\lambda)}) = p_{\mu\lambda}\psi_\lambda,$$

or

$$p_{\mu\lambda} = \frac{\cos(n_{(\mu)}, \overset{*}{n}_{(\lambda)})}{\psi_\lambda}, \tag{2.9}$$

since the cosine of the angle of two vectors $n_{(\mu)}$, $n^*{}_{(\lambda)}$ is defined by

$$\cos(n_{(\mu)}, \overset{*}{n}_{(\lambda)}) = \frac{g(x, n_{(\mu)})n^j_{(\mu)}n^{*i}_{(\lambda)}}{[g_{ij}(x, n_{(\mu)})n^i_{(\mu)}n^j_{(\mu)}]^{\frac{1}{2}}[g_{ij}(x, \overset{*}{n}_{(\lambda)})n^{*i}_{(\lambda)}n^{*j}_{(\lambda)}]^{\frac{1}{2}}},$$

and $n_{(\mu)}$, $n^*{}_{(\lambda)}$ are unit vectors.

We now prove the following theorems.

THEOREM I. *The principal normal of a geodesic G of F_m lies in the space spanned by the secondary normals n^*.*

Proof. We multiply the relation (1.18) by u'^α, obtaining

$$g_{ij}X^j_\gamma \frac{Dx'^i}{Ds} = g_{\alpha\gamma}\frac{\delta u'^\alpha}{\delta s},$$

which is satisfied by the tangent vector u'^α to any curve C in F_m. For a geodesic G we have $\delta u'^\alpha/\delta s = 0$, hence

$$g_{ij}(x, x')X^j_\gamma\left(\frac{Dx'^i}{Ds}\right)_{(G)} = 0.$$

Since the vector Dx'^i/Ds, which defines the principal normal to the geodesics G, satisfies the equation (1.7), it belongs in the space spanned by $n^*{}_{(\mu)}$, therefore

$$\left(\frac{Dx'^i}{Ds}\right)_{(G)} = \sum_{(\mu)} \lambda_{(\mu)}n^{*i}_{(\mu)}, \tag{2.10}$$

where $n^{*i}{}_{(\mu)}$ is a set of $n - m$ orthogonal independent vectors of that space.

THEOREM II. *The tensor $X_{\alpha\beta}{}^i$ considered as a function of a given line element (x^i, x'^i) lies in the space spanned by the secondary normals n^*.*

Proof. We consider the equations

$$X'^i = X^i_\alpha U^\alpha, \qquad g_{ij}X^j_\gamma X^k_\alpha X'^i_{.k} = g_{\beta\gamma}U^\beta_{.\alpha},$$

then we can write

$$g_{ij}X^j_\gamma X^i_\delta P^{*\delta}_{\alpha\beta} = g_{ij}X^j_\gamma\left(\frac{\partial^2 x^i}{\partial u^\alpha \partial u^\beta} + P^{*i}_{hk}X^h_\alpha X^k_\beta\right)$$

and because of (2.3) we obtain

$$g_{ij}(x, x')X^j_\gamma X^i_{\alpha\beta}(u, u') = 0, \tag{2.11}$$

which proves the theorem.

The vector $X_{\alpha\beta}{}^i$ (in i) will be a linear combination of the n^* and therefore

$$X^i_{\alpha\beta} = \sum_{(\mu)} \Omega^*_{(\mu)\alpha\beta}(u, u')n^{*i}_{(\mu)}; \tag{2.12}$$

multiplying the relation (2.12) by $n_{(\nu)}{}^i$ and putting

$$\sum_{(\mu)} \Omega^*_{(\mu)\alpha\beta} \cos (n_{(\nu)}, n^*_{(\mu)}) = \Omega_{(\nu)\alpha\beta},$$

we find

$$n_{(\nu)i} X^i_{\alpha\beta} = \Omega_{(\nu)\alpha\beta}. \tag{2.13}$$

It is obvious that $\Omega_{(\nu)\alpha\beta}$ are tensors symmetric in α, β.

The relations (2.12) and (2.13) are fundamental for the whole theory of subspaces of a Finsler space.

We consider the relation (2.13) and we multiply by $u'^\alpha u'^\beta$, then

$$\Omega_{(\nu)\alpha\beta} u'^\alpha u'^\beta = n_{(\nu)i}\left[\frac{\partial^2 x^i}{\partial u^\alpha \partial u^\beta} u'^\alpha u'^\beta + P^{*i}_{hk} x'^h x'^k\right]; \tag{2.14}$$

but

$$n_{(\nu)} \frac{dx'^i}{ds} = n_{(\nu)i} \frac{\partial^2 x^i}{\partial u^\alpha \partial u^\beta} u'^\alpha u'^\beta. \tag{2.15}$$

Therefore, combining the above equation with (2.14), we obtain

$$\Omega_{(\nu)\alpha\beta} u'^\alpha u'^\beta = n_{(\nu)i}\left[\frac{dx'^i}{ds} + P^{*}_{hk} x'^h x'^k\right] = n_{(\nu)i} \frac{Dx'^i}{Ds}. \tag{2.16}$$

We can easily see that this is the same for all curves of F_m with tangent vector x'^i, but depends on the choice of (x, x'), as in classical differential geometry. Indeed, differentiating the relation

$$n_i x'^i = 0$$

we find

$$\frac{Dn_i}{Ds} x'^i = - n_i \frac{Dx'^i}{Ds},$$

and since $Dn_i/Ds\, x'^i$ depends on x, x' only, so does the right-hand side.

From the identity

$$n_{(\nu)i} \frac{Dx'^i}{Ds} = \left|\frac{Dx'^i}{Ds}\right| \cos\left(n_{(\nu)}, \frac{Dx'}{Ds}\right)$$

we obtain

$$\left|\frac{Dx'^i}{Ds}\right| \equiv \frac{1}{\rho_e} = \frac{\Omega_{(\nu)\alpha\beta} u'^\alpha u'^\beta}{\cos (n_{(\nu)}, Dx'/Ds)}, \tag{2.17}$$

where ρ_e is the radius of curvature of the curve regarded as a curve of F_n. The relation (2.17) may also be written

$$\frac{\cos (n_{(\nu)}, Dx'/Ds)}{\rho_e} = \Omega_{(\nu)\alpha\beta} u'^\alpha u'^\beta, \tag{2.18}$$

and since

$$\Omega_{(\nu)\alpha\beta}u'^{\alpha}u'^{\beta}$$

is the same for all curves of F_m tangent to x'^i, we obtain Meusnier's theorem of classical differential geometry. *We may therefore regard*

$$\Omega_{(\nu)\alpha\beta}u'^{\alpha}u'^{\beta} = \frac{1}{R_{(\nu)}}$$

as the normal curvature corresponding to the normal $n_{(\nu)}{}^i$. It is obvious from (2.17) that the ratio

$$\frac{\Omega_{(\nu)\alpha\beta}u'^{\alpha}u'^{\beta}}{\cos(n_{(\nu)}, Dx'/Ds)}$$

is independent of the choice of $n_{(\nu)}{}^i$.

The concept of the principal direction of a hypersurface F_{n-1} can be extended to any subspace F_m. Indeed, we have shown that to each direction at a point P of F_m correspond $n - m$ normal curvatures

$$(R_{(\nu)}(u, u'))^{-1} = \frac{\Omega_{(\nu)\alpha\beta}(u, u')du^{\alpha}du^{\beta}}{g_{\alpha\beta}(u, u')du^{\alpha}du^{\beta}}$$

associated with the given direction u'.

If we put

$$\Omega_{(\nu)\alpha\beta}(u, u')du^{\alpha}du^{\beta} = 1, \tag{2.19}$$

we obtain a number of $n - m$ loci, of $m - 1$ dimensions each, on the hyperplane spanned by $X_{\alpha}{}^i$, in the Minkowskian tangent space to F_m, at the given point. The principal directions will be given by the extreme values of $g_{\alpha\beta}(u, u')u'^{\alpha}u'^{\beta}$ subject to the conditions (2.19), where u^{α} is kept fixed. In other words, principal directions are directions for which the normal curvatures assume extreme values. According to the multiplier rule we must seek solutions of the equations

$$\frac{\partial}{\partial u'^{\gamma}}\left[g_{\alpha\beta}(u, u')u'^{\alpha}u'^{\beta} + \lambda(\Omega_{(\nu)\alpha\beta}(u, u')u'^{\alpha}u'^{\beta} - 1)\right] = 0,$$

which, after performing the differentiations and using Euler's theorem for homogeneous functions, may be written

$$2\, g_{\alpha\gamma}(u, u')u'^{\alpha} + 2\lambda\Omega_{(\nu)}(u, u')u'^{\alpha} + \lambda\frac{\partial\Omega_{(\nu)}}{\partial u'^{\gamma}}u'^{\alpha}u'^{\beta} = 0. \tag{2.20}$$

The equations (2.20) are of the same type as the corresponding equations for the principal directions of a hypersurface F_{n-1} **(4)**. Applying the same algebraic algorithm, we obtain the following eigenvalue equations:

$$g_{\alpha\gamma}(u, u')u'^{\alpha} = R_{(\nu)}(u, u')\Omega_{(\nu)\alpha\gamma}(u, u')u'^{\alpha}, \tag{2.21}$$

where $(R_{(\nu)}(u, u'))^{-1}$ is the normal curvature corresponding to a solution of (2.21). This is a non-linear eigenvalue problem with eigenvalue $R_{\nu}{}^{-1}$ and little can be said about the number of possible solutions.

Let us assume that at least two independent solutions $u_{(\mu)1}{}'^{\alpha}$, $u_{(\nu)2}{}'^{\alpha}$ corresponding to two distinct normal curvatures $1/R_{(\mu)1}$, $1/R_{(\nu)2}$ exist. Then, from (2.21) we obtain

$$g_{\alpha\gamma}(u, u'_{(\mu)1})u'^{\alpha}_{(\mu)1}u'^{\gamma}_{(\nu)2} = R_{(\mu)1}\Omega_{(\mu)\alpha\gamma}(u, u'_{(\mu)1})u'^{\alpha}_{(\mu)1}u'^{\gamma}_{(\nu)2}$$
$$g_{\alpha\gamma}(u, u'_{(\nu)2})u'^{\alpha}_{(\nu)2}u'^{\gamma}_{(\mu)1} = R_{(\nu)2}\Omega_{(\nu)\alpha\gamma}(u, u'_{(\nu)2})u'^{\alpha}_{(\nu)2}u'^{\gamma}_{(\mu)1},$$

subtracting, we find

$$\frac{\cos(u'_{(\mu)1}, u'_{(\nu)2}) - \cos(u'_{(\mu)2}, u'_{(\nu)1})}{R_{(\mu)1}R_{(\nu)2}} \tag{2.22}$$
$$= \left[\frac{\Omega_{(\mu)\alpha\gamma}(u, u'_{(\mu)1})u'^{\alpha}_{(\mu)1}u'^{\gamma}_{(\nu)2}}{R_{(\nu)2}}\right] - \left[\frac{\Omega_{(\nu)\alpha\gamma}(u, u_{(\nu)2})u'^{\alpha}_{(\nu)2}u'^{\gamma}_{(\mu)1}}{R_{(\mu)1}}\right].$$

When we refer to the same normal $n_{(\nu)}{}^{i}$, the above formula becomes

$$\frac{\cos(u'_{(1)}, u'_{(2)}) - \cos(u'_{(2)}, u'_{(1)})}{R_{(\nu)1}R_{(\nu)2}} \tag{2.23}$$
$$= \left[\frac{\Omega_{(\nu)\alpha\gamma}(u, u'_{(1)})}{R_{(\nu)1}} - \frac{\Omega_{(\nu)\alpha\gamma}(u, u'_{(2)})}{R_{(\nu)2}}\right]u'^{\alpha}_{1}u'^{\gamma}_{2}.$$

The equation (2.23) is a generalization of the orthogonality relation between principal directions of surfaces in classical differential geometry. Indeed, in a locally Euclidean space, the cosine of the angle of two directions is a symmetric function of them. Therefore the left-hand side of (2.23) vanishes and we obtain

$$\frac{\Omega_{(\nu)\alpha\gamma}(u, u'_{(1)})u'^{\alpha}_{1}u'^{\gamma}_{2}}{R_{(\nu)1}} - \frac{\Omega_{(\nu)\alpha\gamma}(u, u'_{(2)})u'^{\alpha}_{1}u'^{\gamma}_{2}}{R_{(\nu)2}} = 0. \tag{2.24}$$

But (2.21) provides

$$\frac{g_{\alpha\gamma}(u)u'^{\alpha}_{1}u'^{\gamma}_{2}}{R_{(\nu)1}} = \Omega_{(\nu)\alpha\gamma}(u, u')u'^{\alpha}_{1}u'^{\gamma}_{2},$$

and thus equation (2.24) becomes

$$\cos(u'_1, u'_2)\left(\frac{1}{R^2_{(\nu)1}} - \frac{1}{R^2_{(\nu)2}}\right) = 0.$$

Since $R_{(\nu)1} \neq R_{(\nu)2}$ we obtain $\cos(u_1', u_2') = 0$, which demonstrates the orthogonality of u_1', u_2'.

We can also define a secondary normal curvature associated to a line element x, x' and depending on Ω^*. For that purpose we consider the relation

$$\frac{Dx'^{i}}{Ds} = \sum \lambda_{(\mu)}n^{*i}_{(\mu)} + X^{i}_{\alpha}\frac{\delta u'^{\alpha}}{\delta s}, \tag{2.25}$$

for an arbitrary curve of F_m and we multiply it by $n_{(\nu)i}$, then

$$n_{(\nu)i}\frac{Dx'^{i}}{Ds} = \sum \lambda_{(\mu)}\cos(n_{(\nu)}, n^{*}_{(\mu)})$$

and

$$\cos\left(n_{(\nu)}, \frac{Dx'}{Ds}\right) = \frac{n_{(\nu)i}\, Dx'^i/Ds}{|Dx'/Ds|} = \rho_e \sum_{\mu} \lambda_{(\mu)} \cos\,(n_{(\nu)}, \overset{*}{n}_{(\mu)}).$$

Or, because of (2.17)

$$\Omega_{(\nu)\alpha\beta} u'^\alpha u'^\beta = \sum_{(\mu)} \lambda_{(\mu)} \cos\,(n_{(\nu)}, \overset{*}{n}_{(\mu)}),$$

and hence, in view of (2.12a),

$$\lambda_{(\mu)} = \overset{*}{\Omega}_{(\mu)\alpha\beta} u'^\alpha u'^\beta.$$

From (2.25) we obtain

$$\frac{Dx'^i}{Ds} = \sum_{(\mu)} (\overset{*}{\Omega}_{(\mu)\alpha\beta} u'^\alpha u'^\beta) \overset{*}{n}{}^i_{(\mu)} + X^i_\alpha \frac{\delta u'^\alpha}{\delta s},$$

and for a geodesic

$$\frac{Dx'^i}{Ds} = \sum_{(\mu)} (\overset{*}{\Omega}_{(\mu)\alpha\beta} u'^\alpha u'^\beta) \overset{*}{n}{}^i_{(\mu)}.$$

We define the secondary normal curvature to be

$$\frac{1}{R^{*2}} = g_{ij}(x, x') \frac{Dx'^i}{Ds} \frac{Dx'^j}{Ds} = \sum_{(\mu)} \psi_{(\mu)}(x, x') \overset{*}{\Omega}_{(\mu)\alpha\beta} \overset{*}{\Omega}_{(\mu)\gamma\delta} u'^\alpha u'^\beta u'^\gamma u'^\delta.$$

In the way $1/R^*$ is defined we see that it is independent of the particular set of normals $n^*_{(\mu)}$.

Let us consider the biquadratic form in the differentials,

$$\phi = \sum_{(\mu)} \psi_{(\mu)}(x, x') \overset{*}{\Omega}_{(\mu)\alpha\beta} \overset{*}{\Omega}_{(\mu)\gamma\delta} du^\alpha du^\beta du^\gamma du^\delta,$$

we may call it the secondary second fundamental form of F_m. Generalizing the concepts of conjugate and asymptotic directions of a surface in classical differential geometry, we may say that two directions at a point defined by du^α and δu^α are conjugate when

$$\sum_{(\mu)} \psi_{(\mu)} \overset{*}{\Omega}_{(\mu)\alpha\beta} \overset{*}{\Omega}_{(\mu)\gamma\delta} du^\alpha \delta u^\beta du^\gamma \delta u^\delta = 0,$$

and asymptotic or self-conjugate when

$$\sum_{(\mu)} \psi_{(\mu)} \overset{*}{\Omega}_{(\mu)\alpha\beta} \overset{*}{\Omega}_{(\mu)\gamma\delta} du^\alpha du^\beta du^\gamma du^\delta = 0.$$

From the above relation and the one defining the secondary normal curvature we conclude that *the secondary normal curvature in an asymptotic direction is always equal to zero* as in Riemannian geometry **(2)**.

3. Covariant derivatives of the normal vectors n^*, n. We define the tensor $n^{*i}{}_{(\mu),\beta}$, covariant derivative of the vector $n^{*i}{}_{(\mu)}$, by projecting $n^{*i}{}_{(\mu),k}$ onto F_m:

$$\overset{*}{n}{}^i_{(\mu),\beta} = \overset{*}{n}{}^i_{(\mu),k} X^k_\beta. \tag{3.1}$$

Obviously

$$n^{*i}_{(\mu),\beta} = \frac{\partial n^{*i}_{(\mu)}}{\partial u^\beta} + P^{*i}_{hk}(x, x') n^{*h}_{(\mu)} X^k_\beta. \tag{3.2}$$

The $n^{*i}{}_{(\mu),\beta}$ are not tangential to F_m, in contrast to Riemannian geometry, and this is the source of much of the difficulty in the derivation of the Gauss-Codazzi equations.

By differentiating the relation (1.7) with respect to u^β and combining the result with (3.2), we obtain

$$\frac{\partial g_{ij}}{\partial u^\beta} n^{*j}_{(\mu)} X^i_\alpha + g_{ij}(x, x') X^i_\alpha (n^{*j}_{(\mu),\beta} - P^{*j}_{hk}(x, x') n^{*h}_{(\mu)} X^k_\beta) + g_{ij} n^{*j}_{(\mu)} \frac{\partial^2 x^i}{\partial u^\alpha \partial u^\beta} = 0,$$

or, after rearranging the terms,

$$g_{ij}(x, x') X^i_\alpha n^{*j}_{(\mu),\beta} + g_{ij} n^{*j}_{(\mu)} \frac{\partial^2 x^i}{\partial u^\alpha \partial u^\beta} + n^{*j}_{(\mu)} X^i_\alpha \left(\frac{\partial g_{ij}}{\partial u^\beta} - g_{hi} P^{*h}_{jk} X^k_\beta \right) = 0.$$

We add and subtract $g_{ij} P^{*i}{}_{hk} X_\alpha{}^h X_\beta{}^k$ in the left-hand side of the above relation, thus obtaining

$$\begin{aligned} g_{ij}(x, x') X^i_\alpha n^{*j}_{(\mu),\beta} + g_{ij} n^{*j}_{(\mu)} \left(\frac{\partial^2 x^i}{\partial u^\alpha \partial u^\beta} + P^{*i}_{hk} X^h_\alpha X^k_\beta \right) \\ + n^{*j}_{(\mu)} X^h_\alpha \left(\frac{\partial g_{hj}}{\partial u^\beta} - g_{lh} P^{*l}_{jk} X^k_\beta - g_{lj} P^{*l}_{hk} X^k_\beta \right) = 0. \end{aligned} \tag{3.3}$$

The term in the last bracket of (3.3) represents the covariant derivative of $g_{ij}\,(x, x')$ with respect to x^k, multiplied by $X_\beta{}^k$; if we put

$$C^*_{hj,k}(x, x') = g_{hj,k}(x, x'),$$

we may write for (3.3),

$$g_{ij}(x, x') X^i_\alpha n^{*j}_{,\beta} + \psi_{(\mu)} \Omega^*_{(\mu)\alpha\beta} + C^*_{ij,k} X^k_\beta X^i_\alpha n^{*j}_{(\mu)} = 0. \tag{3.4}$$

We decompose $n^{*j}{}_{,\beta}$, which is not tangential to F_m, as follows:

$$n^{*j}_{(\mu),\beta} = B^\delta_{(\mu)\beta} X^j_\delta + \sum_{(\kappa)} N^{(\mu)}_{(\kappa)} n^{*j}_{(\kappa)}. \tag{3.5}$$

In order to find $B_{(\mu)\beta}{}^\delta$, we multiply (3.5) by $g_{ij} X_\alpha{}^i$, in view of (1.7) and (3.4), we have

$$B^\epsilon_{(\mu)\beta} = -\psi_{(\mu)} \Omega^*_{(\mu)\alpha\beta} g^{\alpha\epsilon} - C^*_{ihk}(x, x') g^{\alpha\epsilon} X^k_\beta X^i_\alpha n^{*h}_{(\mu)}. \tag{3.6}$$

To obtain the N's we multiply (3.5) by $n^*{}_{(\lambda)j}$, then

$$n^{*i}_{(\mu),\beta} n^*_{(\lambda)i} = N^{(\mu)}_{(\lambda)\beta} \psi_{(\lambda)}. \tag{3.7}$$

The N's are not independent since they satisfy some symmetry conditions which we obtain in the following way. We consider the relations

$$g_{ij}(x, x) n^{*i}_{(\mu)} n^{*j}_{(\lambda)} = \psi_{(\mu)} \delta^\lambda_\mu \qquad \text{(no summation is involved in } \mu\text{)},$$

and we differentiate them with respect to u_β; between the relation which we find and the equation (3.2) we eliminate

$$\frac{\partial n^{*i}_{(\mu)}}{\partial u^\beta},$$

thus

$$(3.8)\quad \frac{\partial g_{ij}(x, x')}{\partial u^\beta} n^{*i}_{(\mu)} n^{*j}_{(\lambda)} + g_{ij}(x, x')[n^{*i}_{(\mu),\beta} n^{*j}_{(\lambda)} + n^{*j}_{(\lambda),\beta} n^{*i}_{(\mu)}$$
$$- P^{*i}_{hk}(x, x') n^{*h}_{(\mu)} n^{*j}_{(\lambda)} X^k_\beta - P^{*j}_{hk}(x, x') n^{*h}_{(\lambda)} n^{*i}_{(\mu)} X^k_\beta] = \frac{\partial \psi_{(\mu)}}{\partial u^\beta} \delta^\lambda_\mu.$$

If we use the relation (3.7), we find

$$(3.9)\qquad \psi_{(\lambda)} N^{(\mu)}_{(\lambda)\beta} + \psi_{(\mu)} N^{(\lambda)}_{(\mu)\beta} = \frac{\partial \psi_{(\mu)}}{\partial u^\beta} \delta^\lambda_\mu - C^*_{ijk} n^{*j}_{(\lambda)} n^{*i}_{(\mu)} X^k_\beta.$$

In conclusion, we have the covariant derivative of $n^*_{(\mu)}$ given by

$$(3.10)\qquad n^{*j}_{(\mu),\beta} = \psi_{(\mu)} X^j_k g^{k\alpha} \Omega_{(\mu)\alpha\beta} - C^*_{ihk} g^{ij} X^k_\beta n^{*h}_{(\mu)} + \sum_{(\lambda)} N^{(\mu)}_{(\lambda)\beta} \psi_{(\lambda)},$$

where the quantities $N_{(\lambda)\beta}{}^{(\mu)}$ (vectors in β) satisfy the symmetry conditions (3.9). In the case of a hypersurface F_{n-1}, the equation (3.10) becomes identical with (4.9) **(4)**, the relation (3.9) giving

$$N_\beta = \frac{1}{\psi} \frac{\partial \psi}{\partial u^\beta} - C_{ijk} n^{*i} n^{*j} X^k_\beta.$$

The equations (3.10) suffer from the disadvantage that the terms $\psi_{(\mu),\beta}$ in the right-hand side involve the derivatives of the tangent x'^i to the curve along which we are differentiating, so that (3.10) depends on the curve under consideration.

As in the case of the $n^{*i}{}_{,\beta}$, we define the covariant derivative $n_{(\mu),\beta}{}^i$ by projecting $n_{(\mu),k}{}^i$ onto F_m,

$$(3.11)\qquad n^i_{(\mu),\beta} = n^i_{(\mu),k} X^k_\beta = \frac{\partial n^i{}_{(\mu)}}{\partial u^\beta} + P^{*i}_{hk}(x, x') n^h_{(\mu)} X^k_\beta,$$

where x'^i is some direction tangential to F_m at the given point. Here we obtain

$$(3.12)\qquad \Omega_{(\mu)\alpha\beta} = - C_{(\mu)\,ijk} X^k_\beta X^i_\alpha n^j_{(\mu)} - g_{ij}(x, n_{(\mu)}) X^i_\alpha n^j_{(\mu),\beta},$$

where

$$(3.13)\qquad C_{(\mu)\,ijk} = g_{ij,k}(x, n_{(\mu)}).$$

We decompose the tensor $n_{(\mu),\beta}{}^j$ as follows:

$$(3.14)\qquad n^i_{(\mu),\beta} = A^\delta_{(\mu)\beta} X^j_\delta + \sum_{(\kappa)} \nu^{(\mu)}_{(\kappa)\beta} n^j_{(\kappa)};$$

multiplying (3.14) by $g_{ij}(x, n_{(\mu)}) X_\alpha{}^i$ we find that

$$(3.15)\qquad A^\delta_{(\mu)\beta} = - \gamma^{\alpha\delta}_{(\mu)} \Omega_{(\mu)\alpha\delta} - \gamma^{\alpha\delta}_{(\mu)} C_{(\mu)\,ijk} X^k_\beta X^i_\alpha n^j_{(\mu)},$$

and therefore

$$(3.16)\quad n^i_{(\mu),\beta} = - \gamma^{\alpha\delta}_{(\mu)} \Omega_{(\mu)\alpha\beta} X^j_\delta - \gamma^{\alpha\delta}_{(\mu)} C_{(\mu)\,ihk} X^j_\delta X^k_\beta X^i_\alpha n^h_{(\mu)} + \sum_{(\kappa)} \nu^{(\mu)}_{(\kappa)\beta} n^j_{(\kappa)}.$$

In order to obtain more information on the ν's we multiply (3.14) by $n_{(\lambda)j}$, then

$$(3.17)\qquad n^j_{(\kappa),\beta}n_{(\lambda)j} = \sum_{(\kappa)} \nu^{(\mu)}_{(\kappa)\beta}n^j_{(\kappa)}n_{(\lambda)j} = \sum_{(\kappa)} \nu^{(\mu)}_{(\kappa)\beta}a_{(\kappa)(\lambda)}$$

where

$$a_{(\kappa)(\lambda)} = n^j_{(\kappa)}n_{(\lambda)j} = \cos(n_{(\lambda)}, n_{(\kappa)}).$$

We note that in general $a_{(\kappa)(\lambda)} \neq a_{(\lambda)(\kappa)}$. Assuming that the determinant $|a_{(\kappa)(\lambda)}|$ is different from zero, we can solve the system (3.17) with respect to the values of $\nu_{(\kappa)}{}_{\beta}{}^{(\mu)}$ and we obtain the ν's as linear combinations of the expressions $n_{(\mu)}{}_{,\beta}{}^{j}n_{(\lambda)j}$, that is,

$$(3.18)\qquad \nu^{(\mu)}_{(\kappa)\beta} = \sum_{(\lambda)} \frac{A^{(\kappa)(\lambda)}}{A}(n^j_{(\mu),\beta}n_{(\lambda)j});$$

where $A^{(\kappa)(\lambda)}$ is the cofactor of the determinant $|a_{(\kappa)(\lambda)}|$ corresponding to the element $a_{(\kappa)(\lambda)}$ and A is the value of that determinant.

As an application of the above theory we may obtain Rodrigues' formula of classical differential geometry.

If we consider the relation (3.16) we may write

$$(3.19)\qquad \frac{Dn^j_{(\mu)}}{Ds} = -\gamma^{\alpha\delta}_{(\mu)}\Omega_{(\mu)\alpha\beta}u'^\beta X^j_\delta - \gamma^{\alpha\delta}_{(\mu)}C_{(\mu)ihk}X^j_\delta X^i_\alpha X'^k n^h_{(\mu)} + \sum_{(\kappa)} \nu^{(\mu)}_{(\kappa)\beta}n^j_{(\kappa)}u'^\beta,$$

since $n_{(\mu)i} = g_{ij}(x, n_{(\mu)})n_{(\mu)}{}^j$, we obtain by differentiation

$$(3.20)\qquad \frac{Dn_{(\mu)i}}{Ds} = C_{(\mu)ijk}n^j_{(\mu)} + g_{ij}(x, n_{(\mu)})\frac{Dn^j_{(\mu)}}{Ds},$$

and substituting (3.19) in (3.20), we find

$$\frac{Dn_{(\mu)i}}{Ds} = -g_{ij}(x, n_{(\mu)})\gamma^{\alpha\delta}_{(\mu)}\Omega_{(\mu)\alpha\beta}u'^\beta X^j_\delta + \sum_{(k)} \nu^{(\mu)}_{(\kappa)\beta}u'^\beta g_{ij}(x, n_{(\mu)})n^j_{(\kappa)}.$$

Multiplying the above equation by $X_\alpha{}^i$, we obtain

$$(3.21)\qquad X^i_\alpha\frac{Dn_{(\mu)i}}{Ds} = -\Omega_{(\mu)\alpha\beta}u'^\beta + \sum_{(k)} \nu^{(\mu)}_{(\kappa)\beta}u'^\beta g_{ij}(x, n_{(\mu)})n^j_{(\kappa)}X^i_\alpha.$$

If $R_{(\mu)}{}^{-1}(x, x')$ is the normal curvature corresponding to a principal direction x'^i of F_m and to a normal $n_{(\mu)}$, we have from equation (2.21)

$$g_{\alpha\beta}(u, u')u'^\beta = R_{(\mu)}(x, x')\Omega_{(\mu)\alpha\beta}u'^\beta.$$

Combining the above relation with (3.21), we write

$$(3.22)\qquad X^i_\alpha\frac{Dn^i_{(\mu)}}{Ds} = -[R_{(\mu)}(x, x')]^{-1}g_{\alpha\beta}(u, u')u'^\beta + \sum_{(k)} \nu^{(\mu)}_{(\kappa)(\beta)}u'^\beta g_{ij}(x, n_{(\mu)})n^j_{(\kappa)}X^i_\alpha.$$

For $m = n - 1$ (hypersurface F_{n-1}), the second term in the right-hand side becomes zero. Indeed in that case the hypersurface has a unique normal n, and therefore, the sum

$$g_{ij}(x, n_{(\mu)})n^j_{(\kappa)}X^i_\alpha$$

is reduced to $g_{ij}(x, n)n^jX_\alpha{}^i$ which is identically equal to zero. But in the case of any subspace F_m, the second term does not vanish, unless we choose a particular set of normals $n_{(\mu)}$ such that

$$\sum_{(\kappa)} \nu^{(\mu)}_{(\kappa)\beta}u'^\beta g_{ij}(x, n_{(\mu)})n^j_{(\kappa)}X^i_\alpha = 0,$$

then

$$X^i_\alpha \frac{Dn^i_{(\mu)}}{Ds} = -\,[R_{(\mu)}(x, x')]^{-1}g_{\alpha\beta}(u, u')u'^\beta,$$

putting $g_{\alpha\beta}(u, u')u'^\beta = y_\alpha$, that is, introducing the covariant component in F_m of u'^β, we find

$$X^i_\alpha \frac{Dn^i_{(\mu)}}{Ds} = -\,(R_{(\mu)}(x, x'))^{-1}y_\alpha. \tag{3.22}$$

The above formula is analogous to Rodrigues' formula and it is similar to the one for a hypersurface **(4)**.

4. The Gauss and Codazzi equations for an F_m. We may obtain relations connecting the curvature tensor of the space F_n with the curvature tensor of F_m and the coefficients Ω^*. To do so, we first consider the covariant derivative of $X_\alpha{}^i$ with respect to u^β (metric in F_m),

$$X^i_{\alpha.\beta} = \frac{\partial X^i_\alpha}{\partial u^\beta} - P^{*\delta}_{\alpha\beta}X^i_\delta.$$

Combining the above relation with relation (2.3), we obtain

$$X^i_{\alpha.\beta} = X^i_{\alpha\beta} - P^{*}_{hk}X^h_\alpha X^k_\beta \tag{4.1}$$

or, because of (2.12),

$$X^i_{\alpha.\beta} = \sum_{(\mu)} \Omega^*_{(\mu)\alpha\beta}n^*_{(\mu)} - P^{*i}_{hk}X^h_\alpha X^k_\beta. \tag{4.2}$$

We know that

$$X^i_{\alpha.\beta\gamma} - X^i_{\alpha.\gamma\beta} = R^\delta_{.\alpha\beta\gamma}X^i_\delta,$$

where $R^\delta{}_{.\alpha\beta\gamma}$ is the curvature tensor of F_m corresponding to the induced connection coefficients $P_{\beta\gamma}{}^{*\alpha}$. By using the expression (4.2), we can write

$$\begin{aligned} R^\delta_{.\alpha\beta\gamma}X^i_\delta = X^h_\alpha&\left[\left(\frac{\partial P^{*i}_{hk}}{\partial u^\gamma} + \frac{\partial P^{*i}_{hk}}{\partial x'^j}\frac{\partial x'^j}{\partial u^\gamma}\right)X^k_\beta - \left(\frac{\partial P^{*i}_{hk}}{\partial u^\beta} + \frac{\partial P^{*i}_{hk}}{\partial x'^j}\frac{\partial x'^j}{\partial u^\beta}\right)X^k_\gamma\right] \\ &- P^{*i}_{hk}(X^h_{\alpha.\gamma}X^k_\beta - X^k_{\alpha.\beta}X^h_\gamma) + \sum_{(\mu)} (\Omega^*_{(\mu)\alpha\beta}n^{*i}_{(\mu).\gamma} - \Omega^*_{(\mu)\alpha\gamma}n^{*i}_{(\mu).\beta}) \\ &+ \sum_{(\mu)} n^{*i}_{(\mu)}(\Omega^*_{(\mu)\alpha\beta.\gamma} - \Omega^*_{(\mu)\alpha\gamma.\beta}) \\ &\qquad\qquad + \sum_{(\mu)} P^{*i}_{hk}n^{*h}_{(\mu)}(\Omega^*_{(\mu)\alpha\gamma}X^k_\beta - \Omega^*_{(\mu)\alpha\beta}X^k_\gamma)\,. \end{aligned} \tag{4.3}$$

With the help of (4.2) the second term of (4.3) can be written

$$X_\alpha^h X_\beta^k X_\gamma^l (P_{kj}^{*i} P_{hl}^{*j} - P_{lj}^{*i} P_{hk}^{*j}) - \sum P_{hk}^{*i} n_{(\mu)}^{*h} (\Omega_{(\mu)\alpha\gamma}^{*} X_\beta^k - \Omega_{(\mu)\alpha\beta}^{*} X_\gamma^k),$$

therefore, (4.3) becomes

$$(4.4)\quad R^{\delta}_{.\alpha\beta\gamma} X_\delta^i = X_\alpha^h \left[\left(\frac{\partial P_{hk}^{*i}}{\partial u^\gamma} + \frac{\partial P_{hk}^{*i}}{\partial x'^j} \frac{\partial x'^j}{\partial u^\gamma} \right) X_\beta^k - \left(\frac{\partial P_{hk}^{*i}}{\partial u^\beta} + \frac{\partial P_{hk}^{*i}}{\partial x'^j} \frac{\partial x'^j}{\partial u^\beta} \right) X_\gamma^k \right]$$
$$+ X_\alpha^h X_\beta^k X_\gamma^l (P_{kj}^{*i} P_{hl}^{*j} - P_{lj}^{*i} P_{hk}^{*j}) + \sum_{(\mu)} (\Omega_{(\mu)\alpha\beta}^{*} n_{(\mu),\gamma}^{*i} - \Omega_{(\mu)\alpha\gamma}^{*} n_{(\mu),\beta}^{*i})$$
$$+ \sum_{(\mu)} n_{(\mu)}^{*i} (\Omega_{(\mu)\alpha\beta,\gamma}^{*} - \Omega_{(\mu)\alpha\gamma,\beta}^{*});$$

the first and the second term in the above equation may be substituted by

$$R^{i}_{.hkl}(x, x') X_\alpha^h X_\beta^k X_\gamma^l$$

according to (A.7), where $R_{hkl}{}^i$ is the curvature tensor of the space F_n. The equation (4.4) then becomes

$$(4.5)\quad X_\delta^i R^{\delta}_{.\alpha\beta\gamma}(u, u') = R^{i}_{.hkl}(x, x') X_\alpha^h X_\beta^k X_\gamma^l + \sum_{(\mu)} (\Omega_{(\mu)\alpha\beta}^{*} n_{(\mu),\gamma}^{*i} - \Omega_{(\mu)\alpha\gamma}^{*} n_{(\mu),\beta}^{*i})$$
$$+ \sum_{(\mu)} n_{(\mu)}^{*i} (\Omega_{(\mu)\alpha\beta,\gamma}^{*} - \Omega_{(\mu)\alpha\gamma,\beta}^{*}).$$

If we use (3.12), we may eliminate the derivatives of $n^*{}_{(\mu)}$ from (4.5) and thus we obtain

$$(4.5a)\quad X_\delta^i R^{\delta}_{.\alpha\beta\gamma}(u, u') = R^{i}_{.hkl}(x, x') X_\alpha^h X_\beta^k X_\gamma^l$$
$$+ \sum_{(\mu)} \psi_{(\mu)} (\Omega_{(\mu)\alpha\beta}^{*} \Omega_{(\mu)\epsilon\gamma}^{*} - \Omega_{(\mu)\alpha\gamma}^{*} \Omega_{(\mu)\epsilon\gamma}^{*}) X_\delta^i g^{\delta\epsilon}$$
$$- C_{jhk}^{*} g^{ij} \sum_{(\mu)} (\Omega_{(\mu)\alpha\beta}^{*} X_\gamma^k - \Omega_{(\mu)\alpha\gamma}^{*} X_\beta^k) n_{(\mu)}^{*h} + \sum_{(\mu)} \sum_{(\lambda)} (N_{(\lambda)\beta}^{(\mu)} \Omega_{(\mu)\alpha\beta}^{*}$$
$$- N_{(\lambda)\gamma}^{(\mu)} \Omega_{(\mu)\alpha\gamma}^{*}) n_{(\lambda)}^{*i} + \sum_{(\mu)} (\Omega_{(\mu)\alpha\beta,\gamma}^{*} - \Omega_{(\mu)\alpha\gamma,\beta}^{*}) n_{(\mu)}^{*i}.$$

Multiplying the above equation by $g_{ij}(x, x') X_\lambda{}^j$, we find

$$(4.6)\quad g_{\delta\lambda} R^{\delta}_{.\alpha\beta\gamma}(u, u') - \sum_{(\mu)} \psi_{(\mu)} (\Omega_{(\mu)\alpha\beta}^{*} \Omega_{(\mu)\lambda\gamma}^{*} - \Omega_{(\mu)\alpha\gamma}^{*} \Omega_{(\mu)\lambda\beta}^{*})$$
$$= g_{ij}(x, x') R^{i}_{.hkl}(x, x') X_\alpha^h X_\beta^k X_\gamma^l X_\lambda^j - X_\lambda^j C_{jhk}^{*} \sum (\Omega_{(\mu)\alpha\beta}^{*} X_\gamma^k - \Omega_{(\mu)\alpha\gamma}^{*} X_\beta^k) n_{(\mu)}^{*h}$$

and multiplying the same equation by $g_{ij}(x, x') n^{*j}{}_{(\nu)}$, we get

$$(4.7)\quad g_{ij}(x, x') n_{(\nu)}^{*j} R^{i}_{.hkl}(x, x') X_\alpha^h X_\beta^k X_\gamma^l - C_{jhk}^{*} \sum_{(\mu)} (\Omega_{(\mu)\alpha\beta}^{*} X_\gamma^k - \Omega_{(\mu)\alpha\gamma}^{*} X_\beta^k) n_{(\mu)}^{*h} n_{(\nu)}^{*j}$$
$$+ \sum_{(\mu)} \psi_{(\mu)} (N_{(\nu)\beta}^{(\mu)} \Omega_{(\mu)\alpha\beta}^{*} - N_{(\nu)\gamma}^{(\mu)} \Omega_{(\mu)\alpha\gamma}^{*}) + \psi_{(\nu)} (\Omega_{(\nu)\alpha\beta,\gamma}^{*} - \Omega_{(\nu)\alpha\gamma,\beta}^{*}) = 0.$$

The equations (4.6) and (4.7) represent a generalization of the Gauss-Codazzi equations of Riemannian geometry.

It is obvious that different forms of Gauss-Codazzi equations are obtained when one considers the fundamental forms $\Omega_{(\nu)\alpha\beta} du^\alpha du^\beta$ together with the normals $n(x)$. For that purpose, we decompose the vector $X_{\alpha\beta}{}^i$ (considered as a vector with respect to the upper index i) into components along the normals n and the tangent plane at the considered point. We put

(4.8) $$X_{\alpha\beta} = \sum_{(\mu)} A_{(\mu)\alpha\beta} n^i_{(\mu)} + W^i_{\alpha\beta}$$

where $W_{\alpha\beta}{}^i$ satisfies the condition

(4.9) $$W^i_{\alpha\beta} n_{(\nu)i} = 0,$$

and by multiplying (4.8) by $n_{(\nu)i}$, we obtain

(4.10) $$\Omega_{(\nu)\alpha\beta} = n_{(\nu)i} X^i_{\alpha\beta} \sum_{(\mu)} A_{(\mu)\alpha\beta} \cos(n_{(\nu)}, n_{(\mu)}),$$

hence $W_{\alpha\beta}{}^i$ is given by the relation

(4.10a) $$W^i_{\alpha\beta} = X^i_{\alpha\beta} - \sum_{(\mu)} A_{(\mu)\alpha\beta} n^i_{(\mu)} = \sum_{(\mu)} \Omega^*_{(\mu)\alpha\beta} n^{*i}_{(\mu)} - \sum_{(\mu)} A_{(\mu)\alpha\beta} n^i_{(\mu)}.$$

Since the vectors n^i are in general different from the vectors n^{*i} and they do not belong in the space spanned by n^{*i}, we look for a decomposition of the n^i's along the n^{*i} and the vectors defining the tangent space to F_m. We decompose the vector n^i in the form

(4.11) $$n^i_{(\mu)} = \sum_{(\lambda)} T^{(\mu)}_{(\lambda)} n^{*i}_{(\lambda)} - M^{\alpha}_{(\mu)} X^i_{\alpha};$$

multiplication by $n_{(\nu)}{}^i$ provides

(4.11a) $$n^i_{(\mu)} n_{(\nu)i} = \sum_{(\lambda)} T^{(\mu)}_{(\lambda)} \cos(n_{(\nu)}, n^*_{(\lambda)});$$

from (4.11) we also obtain

(4.12) $$M^{\delta}_{(\mu)} = n^i_{(\mu)} X^{\delta}_i.$$

Combining (4.11) with (4.10a) and also (4.11a), (4.10), we may write

(4.12a) $$W^i_{\alpha\beta} = \sum_{(\mu)} \Omega^*_{(\mu)\alpha\beta} n^*_{(\mu)} - \sum_{(\lambda)} \sum_{(\mu)} (A_{(\mu)\alpha\beta} T^{(\mu)}_{(\lambda)}) n^*_{(\lambda)} + X^i_{\delta} \sum A_{(\mu)\alpha\beta} M^{\delta}_{(\mu)},$$

(4.13) $$\Omega_{(\nu)\alpha\beta} = \sum_{(\lambda)} \Big(\sum_{(\mu)} A_{(\mu)\alpha\beta} T^{(\mu)}_{(\lambda)} \Big) \cos(n_{(\nu)}, n^*_{(\lambda)}).$$

If we compare the equations (4.13) awith (2.21a), we see that

(4.14) $$\sum_{(\mu)} A_{(\mu)\alpha\beta} T^{(\mu)}_{(\lambda)} = \Omega^*_{(\lambda)\alpha\beta}.$$

In view of the equation (4.14), the relation (4.12a) becomes

(4.15) $$W^i_{\alpha\beta} = X^i_{\delta} \sum_{(\mu)} A_{(\mu)\alpha\beta} M^{\delta}_{(\mu)},$$

thus, $M_{(\mu)}{}^{\delta}$ is given by (4.12), $A_{(\mu)\alpha\beta}$ by (4.10) and $W_{\alpha\beta}{}^i$ by (4.15).

Using the equations (4.1) and (A.8) again, we write

(4.1a) $$X^i_{\alpha,\beta} = \sum_{(\mu)} A_{\alpha\beta} n^i_{(\mu)} + W^i_{\alpha\beta} - P^{*i}_{hk} X^h_{\alpha} X^k_{\beta};$$

differentiating with respect to the metric of F_m and because of (3.11) we obtain

$$(4.16)\quad X^i_{\alpha,\beta\gamma} = \sum_{(\mu)} A_{(\mu)\alpha\beta} n^i_{(\mu),\gamma} + W^i_{\alpha\beta} - \sum_{(\mu)} P^{*\,i}_{hk} A_{(\mu)\alpha\beta} n^h_{(\mu)} X^k_\gamma$$
$$- \left(\frac{\partial P^{*\,i}_{hk}}{\partial u^\gamma} + \frac{\partial P^{*\,i}_{hk}}{\partial x'^j}\frac{\partial x'^j}{\partial u^\gamma}\right) X^h_\alpha X^k_\beta - P^{*\,i}_{hk} X^h_{\alpha,\beta} X^k_\beta - P^{*\,i}_{hk} X^h_\alpha X^k_{\beta,\gamma} + \sum_{(\mu)} A_{(\mu)\alpha\beta,\gamma} n^i_{(\mu)},$$

or

$$(4.17)\quad \sum (A_{(\mu)\alpha\beta,\gamma} - A_{(\mu)\alpha\gamma,\beta}) n^i_{(\mu)} + \sum_{(\mu)} (A_{(\mu)\alpha\beta} n^i_{(\mu),\gamma} - A_{(\mu)\alpha\gamma} n^i_{(\mu),\beta})$$
$$+ X^h_\alpha X^k_\beta X^l_\gamma R^i_{.hkl}(x, x') + W^i_{\alpha\beta,\gamma} - W^i_{\alpha\gamma,\beta} - P^{*\,i}_{hk}(W^h_{\alpha\gamma} X^k_\beta - W^h_{\alpha\beta} X^k_\gamma) = R^\delta_{\alpha\beta\gamma} X^i_\delta.$$

Using the expression for the generalized covariant derivative of $W_{\alpha\beta}{}^i$ with respect to u^γ we find

$$(4.18)\quad X^i_\delta R^\delta_{\alpha\beta\gamma} = \sum_{(\mu)} (A_{(\mu)\alpha\beta,\gamma} - A_{(\mu)\alpha\gamma,\beta}) n^i_{(\mu)} + \sum_{(\mu)} (A_{(\mu)\alpha\beta} n^i_{(\mu),\gamma} - A_{(\mu)\alpha\gamma} n^i_{(\mu),\beta})$$
$$+ X^h_\alpha X^k_\beta X^l_\gamma R^i_{.hkl}(x, x') + W^i_{\alpha\beta\gamma} - W^i_{\alpha\gamma\beta},$$

which, with the help of (3.16), can be written

$$(4.19)\quad X^i_\delta [R^\delta_{\alpha\beta\gamma} - \sum_{(\mu)} \gamma^{\epsilon\delta}_{(\mu)} (\Omega_{(\mu)\epsilon\beta} A_{(\mu)\alpha\gamma} - \Omega_{(\mu)\epsilon\gamma} A_{(\mu)\alpha\beta})] = R^i_{hkl}(x, x') X^h_\alpha X^k_\beta X^l_\gamma$$
$$- \sum_{(\mu)} g^{si}_{(\mu)} C_{(\mu)shk} n^h_{(\mu)} (A_{(\mu)\alpha\beta} X^k_\gamma - A_{(\mu)\alpha\gamma} X^k_\beta) + \sum_{(\mu)} (A_{(\mu)\alpha\beta,\gamma} - A_{(\mu)\alpha\gamma,\beta}) n^i_{(\mu)}$$
$$+ \sum_{(\mu)} \sum_{(\lambda)} n^i_{(\lambda)} (A_{(\mu)\alpha\beta} \nu^{(\mu)}_{(\lambda)\gamma} - A_{(\mu)\alpha\gamma} \nu^{(\lambda)}_{(\mu)\beta}) + W^i_{\alpha\beta\gamma} - W^i_{\alpha\gamma\beta}.$$

The relation (4.19) is important because it provides the Gauss and Codazzi formulae. Indeed, multiplying (4.19) by $g_{ij}(x, n_{(\nu)}) n_{(\nu)}{}^j$ and putting

$$D_{(\mu)(\nu)jhk} = g^{si}(x, n_{(\mu)}) g_{ij}(x, n_{(\nu)}) C_{(\mu)shk},$$
$$m_{(\mu)(\nu)\gamma} = \sum_{(\lambda)} \nu^{(\mu)}_{(\lambda)\gamma} \cos(n_{(\nu)}, n_{(\lambda)}),$$

we obtain the final equation

$$(4.20)\quad \sum_{(\mu)} a_{(\mu)(\nu)} (A_{(\mu)\alpha\beta,\gamma} - A_{(\mu)\alpha\gamma,\beta}) = \sum_{(\mu)} D_{(\mu)(\nu)hkj} n^j_{(\nu)} n^h_{(\mu)} (A_{(\mu)\alpha\beta} X^k_\gamma$$
$$- A_{(\mu)\alpha\gamma} X^k_\beta)$$
$$- \sum_{(\mu)} (m_{(\nu)(\mu)\gamma} A_{(\mu)\alpha\beta} - m_{(\mu)(\nu)\beta} A_{(\mu)\alpha\gamma}) - R^i_{hkl} X^h_\alpha X^k_\beta X^l_\gamma n_{(\nu)i}$$
$$- (W^i_{\alpha\beta\gamma} - W^i_{\alpha\gamma\beta}) g_{ij}(x, n_{(\nu)}) n^j_{(\nu)}.$$

It is possible to remove the terms involving $W_{\alpha\beta\gamma}{}^i$ and replace them by expressions depending on $A_{\alpha\beta}$ or $\Omega_{\alpha\beta}$.

Indeed,

$$(4.21)\quad W^i_{\alpha\beta\gamma} = X^i_\delta \Big(\sum_{(\mu)} A_{(\mu)\alpha\beta\gamma} M^\delta_{(\mu)} + \sum_{(\mu)} A_{(\mu)\alpha\beta} M^\delta_{(\mu),\gamma} \Big)$$
$$+ \Big(\sum_{(\mu)} A_{(\mu)\alpha\beta} M^\delta_{(\mu)} \Big) X^i_{\delta\gamma},$$

and since $X_{\delta\gamma}{}^i = \Sigma_{(\mu)} A_{\delta\gamma} n_{(\mu)}{}^i + W_{\delta\gamma}{}^i$ and $n_{(\nu)i} W_{\delta\gamma}{}^i = 0$, we obtain instead of (4.20),

$$
\begin{aligned}
\sum_{(\mu)} a_{(\mu)(\nu)}(A_{(\mu)\alpha\beta,\gamma} - A_{(\mu)\alpha\gamma,\beta}) = {} & \sum_{(\mu)} D_{(\mu)(\nu)ihk} n^j_{(\mu)} n^h_{(\mu)} (A_{(\mu)\alpha\beta} X^k_\gamma - A_{(\mu)\alpha\gamma} X^k_\beta) \\
& - \sum_{(\mu)} (m_{(\nu)(\mu)\gamma} A_{(\mu)\alpha\beta} - m_{(\nu)(\mu)\beta} A_{(\mu)\alpha\gamma}) \\
& - \sum_{(\mu)} \sum_{(\lambda)} M^\delta_{(\mu)} a_{(\nu)(\lambda)} (A_{(\mu)\alpha\beta} A_{(\lambda)\delta\gamma} - A_{(\mu)\alpha\gamma} A_{(\lambda)\delta\beta}) \\
& - R^i_{hkl} X^h_\alpha X^k_\beta X^l_\gamma n^i_{(\nu)}.
\end{aligned} \tag{4.22}
$$

We consider again the equation (4.19). Multiplying by $g_{ij}(x, n_{(\nu)})X_\xi{}^j$ we obtain

$$
\begin{aligned}
\gamma_{(\nu)\delta\xi}\Big[R^\delta_{\alpha\beta\gamma} - \sum_{(\mu)} \gamma^{\epsilon\delta}_{(\mu)} (\Omega_{(\mu)\epsilon\beta} A_{(\mu)\alpha\gamma} - \Omega_{(\mu)\epsilon\gamma} A_{(\mu)\alpha\beta})\Big] \\
= R^i_{hkl}(x, x') X^h_\alpha X^k_\beta X^l_\gamma X^j_\xi g_{ij}(x, n_{(\nu)}) \\
- \sum_{(\mu)} g^{si}(x, n_{(\mu)}) g_{ij}(x, n_{(\nu)}) C_{(\mu)shk} n^h_{(\mu)} (A_{(\mu)\alpha\beta} X^k_\gamma - A_{(\mu)\alpha\gamma} X^k_\beta) X^j_\xi \\
+ \sum_{(\mu)} (A_{(\mu)\alpha\beta,\gamma} - A_{(\mu)\alpha\gamma,\beta}) g_{ij}(x, n_{(\nu)}) X^j_\xi n^i_{(\mu)} \\
+ \sum_{(\mu)} \sum_{(\lambda)} g_{ij}(x, n_{(\nu)}) X^j_\xi n^i_{(\lambda)} (A_{(\mu)\alpha\beta} \nu_{(\lambda)(\mu)\gamma} - A_{(\mu)\alpha\gamma} \nu_{(\lambda)(\mu)\beta}) \\
+ (W^i_{\alpha\beta\gamma} - W^i_{\alpha\gamma\beta}) g_{ij}(x, n_{(\nu)}) X^j_\xi;
\end{aligned} \tag{4.23}
$$

by eliminating the derivatives $W_{\alpha\beta\gamma}{}^i$ we find a relation

$$
\begin{aligned}
g_{ij}(x, n_{(\nu)}) X^j_\xi (W^i_{\alpha\beta\gamma} - W^i_{\alpha\gamma\beta}) = \gamma_{(\nu)\xi\delta}\Big[\sum_{(\mu)} (A_{(\mu)\alpha\beta,\gamma} - A_{\alpha\gamma,\beta}) M^\delta_{(\mu)} \\
+ \sum_{(\mu)} A_{(\mu)\alpha\beta} M^\delta_{,\gamma} - A_{(\mu)\alpha\gamma} M^\delta_{,\beta}\Big] + g_{ij}(x, n_{(\nu)}) X^j_\xi \Big[\sum_{(\mu)} \sum_{(\lambda)} (A_{(\mu)\alpha\beta} A_{(\lambda)\delta\gamma} \\
- A_{(\mu)\alpha\gamma} A_{(\lambda)\delta\beta}) \,.\, M^\delta_{(\mu)} n^i_{(\lambda)} + X^i_\sigma \sum_{(\mu)} \sum_{(\lambda)} (A_{(\mu)\alpha\beta} A_{(\lambda)\delta\gamma} - A_{(\mu)\alpha\gamma} A_{(\lambda)\delta\beta}) M^\delta_{(\mu)} M^\sigma_{(\lambda)}\Big].
\end{aligned} \tag{4.24}
$$

for the last term of (4.23).

The relations (4.22) *and* (4.23) *thus represent alternative forms of the generalized Gauss and Codazzi equations.*

References

1. H. S. M. Coxeter, *The real projective plane* (Cambridge, 1955).
2. L. P. Eisenhart, *Riemannian Geometry* (Princeton, 1949).
3. H. A. Eliopoulos, *Methods of generalized metric geometry with applications to mathematical physics*, Ph.D. thesis, University of Toronto, August, 1956.
4. H. Rund, *Hypersurfaces of a Finsler space*, Can. J. Math., *8* (1956), 487–503.
5. ——— *Ueber die Parallelverschiebung in Finslerschen Räumen*, Math. Z., *54* (1951), 115–128.
6. ——— *On the analytical properties of curvature tensors in Finsler spaces*, Math. Ann., *127* (1954), 82–104.
7. A. W. Tucker, *On generalized covariant differentiation*, Ann. Math., *32* (1931), 451–60.

Assumption University of Windsor

ON THE IRREDUCIBILITY OF CONVEX BODIES

A. C. WOODS

1. Introduction. We select a Cartesian co-ordinate system in n-dimensional Euclidean space R_n with origin O and employ the usual point-vector notation.

By a lattice Λ in R_n we mean the set of all rational integral combinations of n linearly independent points $X_1, X_2, \ldots, X_n$ of R_n. The points $X_1, X_2, \ldots, X_n$ are said to form a basis of Λ. Let $\{X_1, X_2, \ldots, X_n\}$ denote the determinant formed when the co-ordinates of X_i are taken in order as the ith row of the determinant for $i = 1, 2, \ldots, n$. The absolute value of this determinant is called the determinant $d(\Lambda)$ of Λ. It is well known that $d(\Lambda)$ is independent of the particular basis one takes for Λ.

A star body in R_n is a closed set of points K such that if $X \in K$ then every point of the form tX where $-1 < t < 1$ is an inner point of K. A star body K is called a convex body if it is bounded and satisfies the convex property: if $X \in K$, $Y \in K$ then $tX + (1-t)Y \in K$ provided $0 \leqq t \leqq 1$. It is further called strictly convex if $X \in K$, $Y \in K$ implies that $tX + (1-t)Y$ is an inner point of K when $0 < t < 1$ and $X \neq Y$.

Let Λ be a lattice and K a star body in R_n. We say that Λ is K-admissible if no point of Λ other than 0 is an inner point of K. If K is such that no K-admissible lattice exists then K is said to be of the infinite type, otherwise K is said to be of the finite type. If K is of the finite type the number $\inf d(\Lambda)$ extended over all K-admissible lattices Λ is called the critical determinant $\Delta(K)$ of K and any K-admissible lattice Λ of determinant $d(\Lambda) = \Delta(K)$ is called a critical lattice of K. It is well known that if K is of the finite type then at least one critical lattice of K exists.

Let K be a star body of the finite type in R_n. If K is such that any star body properly contained in K has a smaller critical determinant than K has we say that K is S-irreducible; otherwise K is said to be S-reducible.

Let K be a convex body in R_n. If K is such that any convex body properly contained in K has a smaller critical determinant than K has then we say that K is C-irreducible; otherwise we say that K is C-reducible.

The property of S-irreducibility was first studied by Mahler **(1)** who gave necessary but insufficient conditions for a star body to be S-irreducible. Later **(2)** he considered the property of C-irreducibility and showed that if $n = 2$ then any C-irreducible convex body is also S-irreducible. Rogers **(5)** then gave a set of necessary and sufficient conditions for S-irreducibility which will be stated later.

Received July 3, 1958.

The purpose here is to give an example of a convex body in R_3 that is C-irreducible but not S-irreducible. The proof that the example has these properties relies to a large extent on the work of Whitworth **(6)**. To clarify the picture regarding C-irreducibility we formulate a set of necessary and sufficient conditions for C-rreducibility analogous to the set given by Rogers for S-irreducibility, the proof following similar lines.

2. The set $L(K)$. The results stated in this section are classical.

Let K be a convex body in R_n. We define $L(K)$ to be the set of all points X of the boundary of K such that if X is contained in any line segment of the boundary of K then X is an endpoint of the line segment. Such points are sometimes called extremal points of K so that $L(K)$ constitutes the set of all extremal points of K. As K is symmetric in 0 it is evident that $L(K)$ is also symmetric in 0. Further:

LEMMA 1. *The convex hull of $L(K)$ is K.*

LEMMA 2. *Given $X \in L(K)$ and $\epsilon > 0$ there exists a convex body $K(\epsilon) \subset K$ such that $X \notin K(\epsilon)$ and such that any point of $K - K(\epsilon)$ lies within a distance ϵ of one of the two points $\pm X$.*

3. C-irreducibility. Let K be a star body in R_n. Further let Λ be a critical lattice of K. Let X be a point of Λ on the boundary of K. We say that Λ is free at the point X if, given $\epsilon > 0$, there exists a lattice $\Lambda(\epsilon)$ of determinant $d(\Lambda(\epsilon)) < d(\Lambda) = \Delta(K)$ such that the interior of K contains no point of $\Lambda(\epsilon)$ apart from 0 and any that are within a distance ϵ from one of the two points $\pm X$. Rogers' criterion for S-irreducibility is then as follows:

LEMMA 3. *K is S-irreducible if, and only if, to each point of the boundary of K there corresponds a critical lattice of K that is free at this point.*

We now give an analogous criterion for C-irreducibility.

THEOREM 1. *If K is a convex body then K is C-irreducible if, and only if, to each point of $L(K)$ there corresponds a critical lattice of K that is free at this point.*

Proof. (i) Only if: Assume that K is C-irreducible and let X be an arbitrary point of $L(K)$. By Lemma 2 given $\epsilon > 0$ there exists a convex body $K(\epsilon) \subset K$ such that $X \in K - K(\epsilon)$ and such that any point of $K - K(\epsilon)$ is within a distance ϵ from one of the two points $\pm X$. Since $K(\epsilon)$ is properly contained in K it follows that $\Delta(K(\epsilon)) < \Delta(K)$. Hence there exists a critical lattice $\Lambda(\epsilon)$ of $K(\epsilon)$ of determinant $d(\Lambda(\epsilon)) < d(\Lambda)$. It is evident that K contains no point of $\Lambda(\epsilon)$ in its interior other than 0 and any that may lie within a distance ϵ from one of the two points $\pm X$. Moreover $\Lambda(\epsilon)$ is certainly not K-admissible and therefore taking into account the fact that K is symmetric in 0 we conclude that there must be a point of $\Lambda(\epsilon)$ in the interior of K and

within a distance ϵ from the point X. The sequence $\Lambda(n^{-1})$ of lattices is compact in the sense of Mahler **(3)** and so contains a convergent subsequence with the limit Λ' say. But $\lim_{n\to\infty}K(n^{-1}) = K$ and $\Lambda(n^{-1})$ is a critical lattice of $K(n^{-1})$ for each n, hence Λ' is a critical lattice of K. Further each $\Lambda(n^{-1})$ contains a point within a distance n^{-1} from the point X. Thus Λ' contains X which implies that Λ' is free at X. As X was chosen an arbitrary point of $L(K)$ this proves (i).

(ii) If: Assume that to each point of $L(K)$ there corresponds a critical lattice of K that is free at this point. Take an arbitrary convex body $K' \subset K$ such that $K' \neq K$. There exists a point $X \in L(K) - K'$ for otherwise $L(K) \subset K'$ and so by Lemma 1 $K' = K$ contrary to hypothesis. Let $X \in L(K) - K'$ be fixed. As K' is closed there exists $\epsilon > 0$ such that no point within a distance ϵ from either of the two points $\pm X$ is in K'. By hypothesis there exists a critical lattice Λ of K such that Λ is free at the point X. In particular this implies that there exists a lattice of determinant $d(\Lambda(\epsilon)) < d(\Lambda) = \Delta(K)$ such that no point of $\Lambda(\epsilon)$ apart from 0 and any that may lie within a distance ϵ from one of the two points $\pm X$ is an inner point of K. Hence $\Lambda(\epsilon)$ is K'-admissible from which it follows that $\Delta(K') \leqq d(\Lambda(\epsilon)) < \Delta(K)$. Whence K is C-irreducible. This completes the proof of the theorem.

4. An Example. In looking for a convex body that is C-irreducible and S-reducible we may by Mahler's result confine our attention to dimensions $n \geqq 3$. Further if K is a strictly convex body it is obvious that $L(K)$ is the whole boundary of K. Hence using the previous results K is C-irreducible if, and only if, it is S-irreducible. Again, Dr. Kathleen Ollerenshaw has obtained the following two results **(4)**:

(a) The n-dimensional parallelopiped is S-irreducible for every n.

(b) If K is a two-dimensional S-irreducible convex body then the three-dimensional cylinder on the base K is also S-irreducible.

A more suitable candidate for our purpose has proved to be a sawn-off three-dimensional cube. Whitworth **(6)** has shown that the convex body K in R_3 defined by the inequalities

$$|x_1| \leqq 1, \qquad |x_2| \leqq 1, \qquad |x_3| \leqq 1, \qquad |x_1 + x_2 + x_3| \leqq \tfrac{1}{2}$$

has the critical determinant $\Delta(K) = 3/8$. He has further determined all the critical lattices of K. It is necessary to give a table of these here but before doing so we remark that K has the six automorphisms obtained by permuting the co-ordinates together with the reflections in 0. Thus given any critical lattice of K we obtain six when we apply these transformations. In the following table the only critical lattices of K not included are those obtainable from the ones stated by applying the above automorphisms of K. There are three classes:

Class I: $\Lambda(\rho, \sigma, \beta)$ of basis $X_1 = (\rho - \frac{1}{2}, \sigma - 1, \beta)$, $X_2 = (\rho, \sigma - \frac{1}{2}, \beta - 1)$, $X_3 = (\rho - 1, \sigma, \beta - \frac{1}{2})$ where $\rho + \sigma + \beta = 2$. Another basis for $\Lambda(\rho, \sigma, \beta)$

would be $X_2, X_2 - X_1 = (\frac{1}{2}, \frac{1}{2}, -1)$, $X_3 - X_2 = (-1, \frac{1}{2}, \frac{1}{2})$. The points $X_2 - X_1, X_3 - X_1$ lie in the plane $x_1 + x_2 + x_3 = 0$ while X_2 lies in the plane $x_1 + x_2 + x_3 = \frac{1}{2}$. Hence all points of $\Lambda(\rho, \sigma, \beta)$ that lie on the boundary of K are confined to the three planes $x_1 + x_2 + x_3 = 0$ or $\pm 1/2$. It follows that the same is true of the automorphic images of $\Lambda(\rho, \sigma, \beta)$.

Class II: $\Lambda(\lambda, \mu, \beta)$ of basis $X_1 = (1, -\frac{1}{2}, -\frac{1}{2})$, $X_2 = (-\frac{1}{2}, 1, -\frac{1}{2})$, $X_3 = (-\lambda, -\mu, \beta)$ where $\lambda + \mu - \beta = \frac{1}{2}$, $0 < -\beta \leqq \frac{1}{2}$, $0 \leqq \mu \leqq \frac{1}{2}$, $0 \leqq \lambda \leqq \frac{1}{2}$. The points X_1, X_2 lie in the plane $x_1 + x_2 + x_3 = 0$ while the point X_3 lies in the plane $x_1 + x_2 + x_3 = -\frac{1}{2}$. Hence all points of $\Lambda(\lambda, \mu, \beta)$ that lie on the boundary of K are confined to the three planes $x_1 + x_2 + x_3 = 0$ or $\pm \frac{1}{2}$ and the same is true of the automorphic images of $\Lambda(\lambda, \mu, \beta)$.

Class III: (i) $\Lambda(\nu_1, \nu_2, \chi_1, \chi_2, \beta)$ of basis $X_1 = (-\nu_1, \beta, -\chi_1)$, $X_2 = (-\nu_2, 1 - \beta, -\chi_2)$, $X_3 = (1, -\frac{1}{2}, -\frac{1}{2})$ where $\nu_1 + \nu_2 = \frac{1}{2}$, $\chi_1 + \chi_2 = \frac{1}{2}$, $\beta - \nu_1 - \chi_1 = \pm \frac{1}{2}$. The points X_1, X_2 lie in one of the planes $x_1 + x_2 + x_3 = \pm \frac{1}{2}$ while the point X_3 lies in the plane $x_1 + x_2 + x_3 = 0$ and hence all points of $\Lambda(\nu_1, \nu_2, \chi_1, \chi_2, \beta)$ that are on the boundary of K are confined to the planes $x_1 + x_2 + x_3 = 0$ or $\pm \frac{1}{2}$.

(ii) $\Lambda(\lambda)$ of basis $X_1 = (1, -\frac{1}{2}, -\frac{1}{2})$, $X_2 = (-\lambda, -\frac{1}{2}, 1)$, $X_3 = (\frac{1}{2}, 0, 0)$. Evidently the points of $\Lambda(\lambda)$ that are on the boundary of K are confined to the lines given by $(t, -\frac{1}{2}u_1 - \frac{1}{2}u_2, -\frac{1}{2}u_1 + u_2)$ where u_1, u_2 have one of the following pairs of values: $(0, 0)$, $(1, 0)$, $(-1, 0)$, $(0, 1)$, $(0, -1)$, $(1, 1)$, $(-1, -1)$, $(2, 0)$, $(-2, 0)$. Hence the points of all the automorphic images of $\Lambda(\lambda)$ on the boundary of K are confined to the lines given above together with those obtained from them by permuting the co-ordinates.

(iii) Λ of basis $X_1 = (-\frac{1}{2}, 1, -\frac{1}{2})$, $X_2 = (\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2})$, $X_3 = (\frac{1}{2}, 0, 0)$. The point X_1 lies in the plane $x_1 + x_2 + x_3 = 0$, X_2 in $x_1 + x_2 + x_3 = -\frac{1}{2}$, X_3 in $x_1 + x_2 + x_3 = \frac{1}{2}$; hence all points of Λ that are on the boundary of K are confined to the planes $x_1 + x_2 + x_3 = 0$ or $\pm \frac{1}{2}$. It follows that the same is true of the automorphic images of Λ.

This completes the table of the critical lattices of K. We are now in a position to prove:

THEOREM 2. *K is C-irreducible and S-reducible.*

Proof. We show first that K is S-reducible. From the table given above we see that the only critical lattices of K with points on the boundary of K that do not lie in one of the three planes $x_1 + x_2 + x_3 = 0$ or $\pm \frac{1}{2}$ are those in Class III (ii). The point $(1, -\frac{1}{4}, -\frac{1}{4})$ is on the boundary of K and in the plane $x_1 + x_2 + x_3 = \frac{1}{2}$. Therefore if it is a point of some critical lattice of K it must be in Class III (ii). However, it is obvious that no lattice of this class can contain $(1, -\frac{1}{4}, -\frac{1}{4})$ nor can any lattice which is derived from one of those stated by permuting the co-ordinates. Therefore $(1, -\frac{1}{4}, -\frac{1}{4})$ belongs to no critical lattice of K. By Lemma 3, K is S-reducible.

We now show that K is C-irreducible. The set $L(K)$ consists of the twelve points obtained by permuting the co-ordinates of the point $(1, \frac{1}{2}, -1)$ and taking the six points thus obtained together with their reflections in 0. Hence, by virtue of Theorem 1, K is C-irreducible if we can show that there exists a critical lattice of K which is free at the point $(1, \frac{1}{2}, -1)$. Take the lattice $\Lambda(\frac{1}{2}, 0, 3/2)$ in Class I of the table above. A basis of this lattice is $X_1 = (1, -1, \frac{1}{2})$, $X_2 = (3/2, -\frac{1}{2}, -\frac{1}{2})$, $X_3 = (\frac{1}{2}, 0, 0)$. Another basis would be $Y_1 = X_2 - X_1 = (\frac{1}{2}, \frac{1}{2}, -1)$, $Y_2 = X_1 - X_3 = (\frac{1}{2}, -1, \frac{1}{2})$, $Y_3 = X_3$. The points of $\Lambda(\frac{1}{2}, 0, 3/2)$ on the boundary of K are Y_1, Y_2, Y_3, $Y_1 + Y_2 = (1, -\frac{1}{2}, -\frac{1}{2})$, $Y_1 + Y_3 = (1, \frac{1}{2}, -1)$, $Y_1 - Y_3 = (0, \frac{1}{2}, -1)$, $Y_2 - Y_3 = (0, -1, \frac{1}{2})$, $Y_2 + Y_3 = (1, -1, \frac{1}{2})$, $Y_1 + Y_2 - Y_3 = (\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2})$ together with their reflections in 0. In particular we see that $Y_1 + Y_3 = (1, \frac{1}{2}, -1)$ is a point of the lattice. For a given $\delta > 0$ denote by $\Lambda(\delta)$ the lattice of basis $Y_1' = (\frac{1}{2} - \delta, \frac{1}{2}, -1)$, $Y_2' = (\frac{1}{2} + \delta, -1, \frac{1}{2})$, $Y_3' = (\frac{1}{2} - \delta, 0, \delta)$. Evidently as $\delta \to 0$ so $Y_1' \to Y_1$, $Y_2' \to Y_2$, $Y_3' \to Y_3$ and therefore also $\Lambda(\delta) \to \Lambda(\frac{1}{2}, 0, 3/2)$. Moreover,

$$d(\Lambda(\delta)) = \left\| \begin{matrix} \frac{1}{2} - \delta & \frac{1}{2} & -1 \\ \frac{1}{2} + \delta & -1 & \frac{1}{2} \\ \frac{1}{2} - \delta & 0 & \delta \end{matrix} \right\| = \frac{3}{8} - \frac{1}{2}\delta^2 < \frac{3}{8}$$

provided only that δ is sufficiently small. Since in the limit $\delta \to 0$ the basis given for $\Lambda(\delta)$ becomes the basis given for $\Lambda(\frac{1}{2}, 0, 3/2)$ it follows that for all sufficiently small δ the only points of $\Lambda(\delta)$ that can lie in the interior of K are

$$Y_1' = (\tfrac{1}{2} - \delta, \tfrac{1}{2}, -1),\ Y_2'\ (\tfrac{1}{2} + \delta, -1, \tfrac{1}{2}),\ Y_3' = (\tfrac{1}{2} - \delta, 0, \delta),$$

$$Y_1' + Y_2' = (1, -\tfrac{1}{2}, -\tfrac{1}{2}),\ Y_1' + Y_3' = (1 - 2\delta, \tfrac{1}{2}, \delta - 1),\ Y_2' + Y_3' = (1, -1, \tfrac{1}{2} + \delta),$$

$$Y_1' - Y_3' = (0, \tfrac{1}{2}, -1 - \delta),\ Y_2' - Y_3' = (2\delta, -1, \tfrac{1}{2} - \delta),\ Y_1' + Y_2' - Y_3' = (\tfrac{1}{2} + \delta, -\tfrac{1}{2}, -\tfrac{1}{2} - \delta)$$

together with their reflections in 0. But it is clear that the only ones in the interior of K are $\pm (Y_1' + Y_3')$. Moreover

$$\lim_{\delta \to R} (Y_1' + Y_3') = (1, \tfrac{1}{2}, -1),$$

hence $\Lambda(3/2, 0, \frac{1}{2})$ is free at the point $(1, \frac{1}{2}, -1)$. Therefore K is C-irreducible. This completes the proof of Theorem 2.

Part of this work is extracted from a thesis for the degree of Doctor of Philosophy at the University of Manchester, written under the supervision of Professor K. Mahler to whom I am very grateful for advice and encouragement.

References

1. K. Mahler, *Lattice points in n-dimensional star bodies II, Reducibility theorems*, Proc. Nederl. Akad. Wetensch., *49* (1946), 331–43.
2. ——— *On irreducible convex domains*, Proc. Nederl. Akad. Wetensch., *50* (1947), 98–107.
3. ——— *Lattice points in n-dimensional star bodies, I, Existence theorems*, Proc. Roy. Soc. London, Ser. A, *187* (1946), 151–87.
4. K. Ollerenshaw, *Irreducible convex bodies*, Quart. J. Math., Oxford (2), *4* (1953), 293–302.
5. C. A. Rogers, *A note on irreducible star bodies*, Proc. Nederl. Akad. Wetensch., *50* (1947), 868–72.
6. J. V. Whitworth, *On the densest packing of sections of a cube*, Ann. Mat. Pura Appl., Ser. 4, *27* (1948), 29–37.

Tulane University of Louisiana

DENSE SUBGRAPHS AND CONNECTIVITY

R. E. NETTLETON, K. GOLDBERG, AND M. S. GREEN

A proper subgraph of a connected linear graph is said to disconnect the graph if removing it leaves a disconnected graph. In this paper we characterize, in the following sense, the disconnecting subgraphs of a fixed connected graph. We define two distinct types of disconnecting subgraphs (isthmuses and articulators) which are minimal in the sense that no proper subgraph of either type can disconnect the graph. We then show that any disconnecting subgraph must contain either an isthmus or an articulator. We also define a set of subgraphs (called dense) which form a lattice. We show that the union of the minimal dense subgraphs contains all isthmuses and articulators. In terms of these subgraphs we investigate some of the consequences of assuming that a disconnecting subgraph must contain at least m points.

1. Definitions. A (linear, undirected) graph G is a finite set of elements $p_1, p_2, \ldots, p_n$ called *points*, and a set of ordered pairs of these elements defining a symmetric, non-reflexive binary relation. Two points occurring in an ordered pair are said to be *neighbours*. A *subgraph* of G is a subset of the points of G together with all the ordered pairs in G containing only elements of the subset. A subgraph is thus determined by its set of points when the binary relation of G is understood.

Two distinct points, p and q, in G are said to be *connected* by a *path* of *length* k if there exist $k + 1$ distinct points $p = p_1, p_2, \ldots, p_{k+1} = q$ such that the ordered pairs (p_i, p_{i+1}), for $i = 1, 2, \ldots, k$, are in G. The *distance* between two points is the length of the shortest path between them. The *diameter* of the graph is the greatest distance between pairs of points in the graph.

A graph having only one point or more than one point and every pair of points connected is also called *connected*. If every pair of points are neighbours the graph is called *completely connected*. A graph which is not connected is called *disconnected*. The null graph is disconnected.

The *union*, *intersection*, and *difference* of two subgraphs G_1 and G_2 is the subgraph whose set of points is the union, intersection, or difference of the sets of points of G_1 and G_2. We denote the union by $G_1 + G_2$ and the difference by $G_1 - G_2$.

If a graph is not connected it is the union of a set of disjoint subgraphs each one of which is connected and such that the union of any two is not connected. This set is unique and we refer to it as the *partition* of the graph.

Received March 27, 1958.

We say that a proper subgraph G' of a connected graph G *disconnects* G if $G - G'$ is disconnected. We shall be interested in ways of disconnecting a fixed connected graph G containing n points and to this end we introduce two definitions.

A *k-isthmus* of G is a completely connected subgraph which has k points, disconnects G, but does not properly contain a completely connected subgraph which disconnects G. A *k-articulator* of G is a subgraph G' which has k points, disconnects G, is not completely connected, and has the property that each subgraph in the partition of $G - G'$ has a neighbour of each point in G'. We shall use the generic terms *isthmus* and *articulator* when the number of points is irrelevant.

For example, if we denote G pictorially with lines representing the relation between points we can see the isthmuses and articulators in the following connected graphs:

FIGURE 1

FIGURE 2

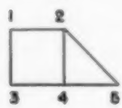

FIGURE 3

In Figure 1 the subgraphs with point sets $\{1, 4\}$ and $\{2, 3\}$ are articulators but there are no isthmuses. In Figure 2 $\{2, 3\}$ is an isthmus but there are no articulators. In Figure 3 $\{2, 4\}$ is an isthmus and $\{1, 4\}$ and $\{2, 3\}$ are articulators.

We now define a type of subgraph which we shall prove has a close connection with the isthmuses and articulators of G. A connected subgraph G' is called *dense* if $G' = G$ or if every point in $G - G'$ has a neighbour in G'. A dense proper subgraph which is contained in no other dense subgraph except G we call *D-maximal*; a dense subgraph containing no other dense subgraph we call *D-minimal*. We let S_D denote the collection of dense subgraphs of G ordered by inclusion together with the empty graph ϕ. We let Γ denote the union of all D-minimal subgraphs of G. Unless otherwise stated all dense subgraphs, isthmuses, and articulators are those of G.

We call G *m-connected* if $G - G'$ is connected for every subgraph G' containing fewer than m points.

The subgraph of neighbours of a point p in a subgraph G' we denote by $G'(p)$.

2. Dense Subgraphs. Suppose G_1 is a dense subgraph and G_2 is a subgraph containing G_1. Since every point in $G_2 - G_1$ has a neighbour in the connected graph G_1 it follows that G_2 is connected. Also every point in $G - G_2$ has a neighbour in G_2 (in fact in G_1). Therefore we have

LEMMA 2.1. *A subgraph which contains a dense subgraph is also dense.*

Let G_1 and G_2 be two dense subgraphs. By this lemma their union is also dense. Their intersection need not be dense but if it is not it cannot contain a dense subgraph, again by this lemma. Therefore we have

THEOREM 2.2. *S_D is a lattice in which the* l.u.b. *is the graph union, and the* g.l.b. *is the graph intersection when the intersection is dense and otherwise is ϕ.*

Applying Lemma 2.1 to the definition of D-maximal we get

LEMMA 2.3. *A subgraph is D-maximal if and only if it is connected and contains $n - 1$ points.*

Suppose $n > 2$ and let d denote the diameter of G. Let p_1 and p_2 be points such that the distance between them is d. If $G - p_1$ is not connected let $\{G_i\}$ denote its partition. Suppose p_2 is in G_1 and let p_3 be a point in G_2 which is a neighbour of p_1. Any path between p_2 and p_3 must pass through p_1 since removing p_1 disconnects an otherwise connected graph. Thus the distance between p_2 and p_3 is $d + 1$ which is a contradiction. It follows that $G - p_1$, and $G - p_2$ by symmetry, is connected and so a D-maximal subgraph. We have proved

THEOREM 2.4. *If $n > 2$ then G contains at least two D-maximal subgraphs. Thus every point is contained in a proper dense subgraph.*

We shall show that if $G = \Gamma$ and the D-minimal subgraphs are mutually disjoint then G is completely connected. First we need

LEMMA 2.5. *Let $\{\Gamma_i\}$ be a collection of mutually disjoint dense subgraphs whose union is G. If at least one of the Γ_i contains more than one point then there is a dense subgraph containing none of the Γ_i.*

We shall prove this by constructing the desired dense subgraph G'.

Let $\Gamma_1, \Gamma_2, \ldots, \Gamma_s$ be those subgraphs among the Γ_i containing more than one point.

If $s = 1$ let G' denote an arbitrary D-maximal subgraph of Γ_1. Every point in Γ_1 has a neighbour in G' and every other point in G is dense and so is a neighbour to every point in G'. Thus G' is a dense subgraph of G properly contained in Γ_1 and disjoint from the other Γ_i, as desired.

Now suppose $s > 2$. Choose an arbitrary point q in Γ_1. Let $q_1 = q$ and q_i be a neighbour of q in Γ_i for $i = 2, 3, \ldots, s$. Since each Γ_i $(i = 1, 2, \ldots, s)$ contains more than one point, it contains a D-maximal subgraph G_i which contains q_i. Let G' be the union of the G_i.

Each of the G_i is connected and each has a neighbour of q or contains q so G' is connected. Let p be an arbitrary point in G distinct from q. Let Γ_i be that subgraph containing p. If Γ_i contains no other points p is a neighbour of every point in G'. If Γ_i has more than one point then either p has a neighbour in G_i (if p is not in G_i or G_i contains more than one point) or is a neigh-

bour of q (if p is the only point in G_i). Thus G' is dense and we can complete our argument as before.

We can now prove

THEOREM 2.6. *If $G = \Gamma$ and the D-minimal subgraphs are mutually disjoint then G is completely connected.*

Given the hypothesis, if any D-minimal subgraph contains more than one point we can apply Lemma 2.5 to obtain a dense subgraph not containing any D-minimal subgraph. Since this is absurd every D-minimal subgraph contains exactly one point. Therefore every point of G is dense and so G is completely connected.

Now we prove

THEOREM 2.7. *If a point p is not D-minimal then $\Gamma(p)$ disconnects G.*

Suppose $G - \Gamma(p)$ is connected. It contains p and every point in $\Gamma(p)$ is a neighbour of p so that $G - \Gamma(p)$ is dense. Thus $G - \Gamma(p)$ contains a D-minimal subgraph G' having no neighbours of p. This is possible only if $G' = p$.

We have incidentally proved

LEMMA 2.8. *If $G - (G(p) - H)$ is connected it is dense.*

LEMMA 2.9. *$G - G(p)$ is connected if and only if p is D-minimal.*

3. Connectivity. We begin by finding a necessary and sufficient condition that a subgraph be an articulator or an isthmus.

If G' is an articulator or an isthmus then $G - G'$ is not connected. Let p be any point in G' and consider $G - G' + p$. If G' is an articulator every subgraph in the partition of $G - G'$ contains a neighbour of p so $G - G' + p$ is connected. Likewise every point in $G' - p$ has a neighbour in $G - G'$ and so in $G - G' + p = G - (G' - p)$. Therefore $G - G' + p$ is dense as is $G - G''$ for every proper subgraph G'' of G'. If G' is an isthmus then $G' - p$ is completely connected so $G - G' + p$ is connected. Every point in $G' - p$ is a neighbour of p so $G - G' + p$ is dense as is $G - G''$ for every proper subgraph G'' of G'.

Now suppose G' is a subgraph which disconnects G but $G - G' + p$ is connected for every point p in G'. Then such a point must have a neighbour in every subgraph of the partition of $G - G'$ so G' is an articulator if it is not completely connected and an isthmus if it is. Thus we have

THEOREM 3.1. *A subgraph G' is either an articulator or an isthmus if and only if it disconnects G and $G - G''$ is connected* (*and so dense*) *for every proper subgraph G'' of G'.*

COROLLARY. *An articulator does not properly contain an articulator. An articulator does not contain an isthmus and conversely.*

Let G' be an articulator or isthmus and let p be any point in G'. Then, by Theorem 3.1, $G - G' + p$ is dense and so contains a D-minimal subgraph which must contain p since $G - G'$ is not dense. It follows that every point in G' is in Γ. That is

THEOREM 3.2. *All articulators and all isthmuses are contained in* Γ.

Let G' be any subgraph which disconnects G. It is either an articulator or an isthmus or, by Theorem 3.1, contains a proper subgraph G'' which disconnects G. By repeating the argument on G'' we are eventually led to the case when $G - p$ is not connected for a point p. Since such a point p is an isthmus we have

THEOREM 3.3. *A subgraph which disconnects G contains an articulator or an isthmus.*

We now turn to some of the consequences of m-connectivity and obtain

THEOREM 3.4. *If G is m-connected then*

1. $G - G'$ *is dense for every subgraph* G' *containing less than* m *points, and conversely.*

2. G *contains no* k*-isthmus for* $k = 1, 2, \ldots, m - 1$ *and no* k*-articulator for* $k = 2, 3, \ldots, m - 1$, *and conversely.*

3. Γ *contains at least* m *points as does* $\Gamma(p)$ *for every point* p *in* G *which is not D-minimal.*

4. *The intersection of any* $m - 1$ *D-maximal subgraphs is dense.*

5. *An m-articulator of* Γ *is an m-articulator of* G.

6. *Either* G *is completely connected and* $n = m$, *or it is not and* $n > m + 2$.

7. *If* p *is a point in* $G - \Gamma$ *and* q *is a point in* $G(p)$ *then* $G(q)$ *has more than* m *points.*

Let G be m-connected and G' be a subgraph containing less than m points. Suppose there is a point p in G' without a neighbour in $G - G'$. Then $G - G' + p = G - (G' - p)$ is not connected contrary to the definition of m-connected. It follows that every point p in G' has a neighbour in the connected subgraph $G - G'$ which is thus dense. The converse is clear and so part 1 is proved.

The necessity of part 2 follows from Theorem 3.3 and the sufficiency is clear.

Since the second half of part 3 follows from Theorem 2.7 we must show that Γ contains at least m points. If G is completely connected then $G = \Gamma$ so Γ is m-connected and thus contains at least m points. If G is not completely connected it contains at least one point which is not D-minimal. That point has at least m neighbours in Γ so Γ contains at least m points.

Part 4 follows from part 1 and Lemma 2.3.

Part 5 follows from

LEMMA 3.5. *An articulator of* Γ *disconnects* G.

In proving this we do not assume that G is m-connected.

Let Γ' be an articulator of Γ and suppose $G - \Gamma'$ is connected. Every point in Γ' has a neighbour in $\Gamma - \Gamma'$ and so in $G - \Gamma'$. Therefore $G - \Gamma'$ is dense and so contains a D-minimal subgraph G'. But $G - \Gamma' - (G - \Gamma) = \Gamma - \Gamma'$ is not dense. This implies that $G - \Gamma$ contains some point of G' contrary to the definition of Γ. It follows that Γ' disconnects G.

Now suppose G is m-connected and Γ' contains m points. By this lemma and Theorem 3.3, Γ' contains either an articulator or an isthmus. But it cannot contain either properly by part 2. Since Γ' is not completely connected it is an articulator (of G).

If G is completely connected then $n = m$. Otherwise there are points p and q in G which are not connected so $G - p - q$ disconnects G. It follows that $m < n - 2$ and part 6 is proved.

If p is a point in $G - \Gamma$ and q is a point in $G(p)$ then if q is not D-minimal it has at least m neighbours in Γ as well as at least one (that is p) in $G - \Gamma$. If q is D-minimal then it has $n - 1$ neighbours. Since $G - \Gamma$ has a point G cannot be completely connected so $n - 1 > m$ and part 7 is proved.

As a partial converse of part 4 we prove

THEOREM 3.6. *If the intersection of any* $m > 2$ *D-maximal subgraphs is connected then there are no k-isthmuses or k-articulators for* $m > k > 2$.

Let G' be a k-isthmus or k-articulator with $m > k > 2$, and let p be an arbitrary point in G'. By Theorem 3.1 we know that $G - p$ is dense and so D-maximal. Thus $G - G'$ is the intersection of $k < m$ D-maximal subgraphs but is not connected contrary to the hypothesis of the theorem.

We complete this section with a few isolated results.

THEOREM 3.7. *If* G *is not completely connected but* $G(p)$ *is for some point* p *then* p *is in* $G - \Gamma$.

Suppose Γ' is a D-minimal subgraph containing p. If $\Gamma' = p$ then $G(p) = G - p$ so G is completely connected contrary to hypothesis. Therefore Γ' contains a neighbour q of p. Since every point which is a neighbour of p is a neighbour of q we see that $\Gamma' - p$ is dense, again contrary to hypothesis. Thus p is not in any D-minimal subgraph.

THEOREM 3.8. *The intersection of all dense subgraphs is exactly the subgraph of* 1*-isthmuses.*

Suppose p is a point contained in all dense subgraphs. Then $G - p$ does not contain a dense subgraph and so is not dense. Since p has a neighbour in $G - p$ the latter is not connected so p is a 1-isthmus. Conversely, if p is a 1-isthmus $G - p$ is not dense but G is so that every dense subgraph contains p.

Added in proof: In order to complete the statements of Theorems 3.1 and 3.3 we should have proved that *if G is not completely connected it contains a disconnecting subgraph*. This is trivial. For suppose G is not completely connected. Then it contains at least three points, at least two of which are not neighbours. Then $G - p - q$ disconnects G.

The authors are indebted to A. J. Hoffman of the General Electric Company for his many helpul suggestions.

References

1. F. Harary and R. Z. Norman, *The dissimilarity characteristic of Husimi trees*, Ann. Math. *58* (1953), pp. 134–41.
2. D. König, *Theorie der endlichen und unendlichen Graphen* (Leipzig, 1936).

National Bureau of Standards
Washington, D.C.

THE TERM AND STOCHASTIC RANKS OF A MATRIX

A. L. DULMAGE AND N. S. MENDELSOHN

1. Introduction. The *term rank* ρ of a matrix is the order of the largest minor which has a non-zero term in the expansion of its determinant. In a recent paper **(1)**, the authors made the following conjecture. If S is the sum of all the entries in a square matrix of non-negative real numbers and if M is the maximum row or column sum, then the term rank ρ of the matrix is greater than or equal to the least integer which is greater than or equal to S/M. A generalization of this conjecture is proved in § 2.

The term *doubly stochastic* has been used to describe a matrix of non-negative entries in which the row and column sums are all equal to one. In this paper, by a doubly stochastic matrix, the authors mean a matrix of non-negative entries in which the row and column sums are all equal to the same real number T. If an $n \times n$ matrix A is embedded by the addition to A of r rows and columns in an $(n + r) \times (n + r)$ matrix B with row and column sums equal to T, we say that B is an (r, T) *doubly stochastic* (*abbreviated as* (r, T) *d.s.*) *extension of* A. In **(1)**, the authors made use of a d.s. extension of a matrix A to obtain an estimate of the term rank of A. In this paper, the authors describe all such extensions and give a necessary and sufficient condition that a matrix B be a vertex matrix of the convex set of all (r, T) d.s. extensions of A.

For a square matrix of non-negative entries, the concept of *stochastic rank* is introduced. Some results concerning this rank are obtained and the connection between it and term rank is noted.

In the final section, the problem of finding all d.s. extensions of a matrix A is formulated as a linear programming problem.

2. A lower bound for term rank. Let I and J be arbitrary sets and let $f(i, j)$, $i \in I, j \in J$, be a real-valued non-negative function on $I \times J$ which is not identically zero. The concept of term rank can be extended to such a function $f(i, j)$ as follows. A finite set of pairs (i_1, j_1), $(i_2, j_2), \ldots, (i_r, j_r)$ is *disjoint* if $i_p = i_q$ only if $p = q$ and if $j_p = j_q$ only if $p = q$. A function $f(i, j)$ has term rank ρ if, and only if, there exists a disjoint set of pairs (i_1, j_1), $(i_2, j_2), \ldots, (i_\rho, j_\rho)$ such that $f(i_r, j_r) > 0$ for $r = 1, 2, \ldots, \rho$ but for any disjoint set consisting of $\rho + 1$ pairs, $f(i, j) = 0$ for at least one pair of the set. If no such maximal disjoint set exists, the term rank is infinite.

Let σ be the collection of finite subsets of I and τ the collection of finite subsets of J. In this setting, we have the following theorem.

Received May 8, 1958.

THEOREM 1. *If $f(i,j)$ satisfies the conditions*

(i) $$R_i = \sup_{B\in\tau}\Big[\sum_{j\in B} f(i,j)\Big] \quad \textit{is finite for all } i \in I$$

and

(ii) $$C_j = \sup_{A\in\sigma}\Big[\sum_{i\in A} f(i,j)\Big] \quad \textit{is finite for all } j \in J$$

then either the term rank ρ of the function $f(i,j)$ is infinite or

$$S = \sup_{\substack{A\in\sigma\\ B\in\tau}}\Big[\sum_{i\in A}\sum_{j\in B} f(i,j)\Big] \textit{ and } M = \sup_{\substack{i\in I\\ j\in J}}[R_i, C_j]$$

are finite and ρ is greater than or equal to the least integer which is greater than or equal to S/M.

Proof. Let K be the graph of which the edges are the pairs (i,j) for which $f(i,j) > 0$. The vertex sets of this (bipartite) graph K are I and J. If ρ is finite, the exterior dimension (see **(3)**) of K is equal to ρ. If $[P,Q]$ is a minimal exterior pair for K and if U, V is any pair of finite subsets of I and J, then, since $f(i,j) = 0$ for $i \in \bar{P}$ and $j \in \bar{Q}$, we have

$$\begin{aligned}\sum_{i\in U}\sum_{j\in V} f(i,j) &= \sum_{i\in U\cap P}\sum_{j\in V\cap Q} f(i,j)\\ &\quad + \sum_{i\in U\cap \bar{P}}\sum_{j\in V\cap Q} f(i,j) + \sum_{i\in U\cap P}\sum_{j\in V\cap \bar{Q}} f(i,j)\\ &\le \sum_{i\in U}\sum_{j\in V\cap Q} f(i,j) + \sum_{i\in U\cap P}\sum_{j\in V} f(i,j)\\ &\le \sum_{i\in P} R_i + \sum_{j\in Q} C_j\end{aligned}$$

which is finite and independent of U and V.

Now

$$S = \sup_{\substack{U\in\sigma\\ V\in\tau}}\sum_{i\in U}\sum_{j\in V} f(i,j) \le \sum_{i\in P} R_i + \sum_{j\in Q} C_j$$

so that S is finite. Further

$$\begin{aligned}R_i &= \sup_{V\in\sigma}\Big[\sum_{j\in V} f(i,j)\Big]\\ &\le \sup_{\substack{U\in\sigma\\ V\in\tau}}\Big[\sum_{i\in U}\sum_{j\in V} f(i,j)\Big] = S,\end{aligned}$$

for all i. Similarly $C_j \le S$ for all j. Thus,

$$M = \sup_{\substack{i\in I\\ j\in J}}[R_i, C_j] \le S \quad \text{so that } M \text{ is finite.}$$

Now, let t be the unique integer such that $t - 1 < S/M \le t$. We must show $\rho \ge t$. If $\rho < t$ then, since ρ is integral, we have $\rho \le t - 1$. If $[P,Q]$ is

a minimal exterior pair for K and $\nu(P)$ denotes the number of elements in P then $\rho = \nu(P) + \nu(Q)$ (**3**, Theorem 2). It follows that

$$\rho M = (\nu(P) + \nu(Q))M \geq \sum_{i \in P} R_i + \sum_{j \in Q} C_j \geq S.$$

Thus $S/M \leq \rho \leq t - 1$, a contradiction.

If the sets I and J in Theorem 1 are finite sets of orders n and m, ρ becomes the term rank of an $n \times m$ matrix a_{ij} in which $a_{ij} = f(i, j)$, M becomes the maximum row or column sum and S is the sum of all the entries in the matrix. If, in addition, $n = m$, Theorem 1 reduces to the conjecture in (**1**) referred to in the Introduction.

3. The stochastic rank of a matrix. Let A be an $n \times n$ matrix with non-negative entries a_{ij}. If M is the maximum row or column sum in A, then, for every $T \geq M$, and for every integer $r \geq n$, there exists a matrix B which is an (r, T) d.s. extension of A. In fact, if

$$R_i = \sum_{j=1}^{n} a_{ij}$$

and

$$C_j = \sum_{i=1}^{n} a_{ij}$$

for $i, j = 1, 2, \ldots, n$, the matrix $B = (b_{ij})$ may be defined as follows

$$\begin{aligned}
b_{ij} &= a_{ij} \text{ for } i \leq n,\ j \leq n,\\
b_{ij} &= a_{2n+1-j,\,2n+1-i}, && \text{for } n+1 \leq i \leq 2n, \quad n+1 \leq j \leq 2n.\\
b_{ij} &= 0 && \text{for } i \leq n, n+1 \leq j \leq 2n \text{ provided } i+j \neq 2n+1,\\
&= T - R_i; && \text{for } i+j = 2n+1,\ i \leq n,\\
b_{ij} &= 0 && \text{for } n+1 \leq i \leq 2n,\ j \leq n \text{ provided } i+j \neq 2n+1,\\
&= T - C_j && \text{for } i+j = 2n+1,\ j \leq n,\\
b_{ij} &= 0 && \text{for } 2n+1 \leq i \leq n+r, \text{ or } 2n+1 \leq j \leq n+r, \text{ provided } i \neq j\\
&= T && \text{for } 2n+1 \leq i = j \leq n+r.
\end{aligned}$$

The question naturally arises, for what $r \leq n - 1$ and $T \geq M$ is an (r, T) d.s. extension of A possible? In Theorem 2, we have a complete answer to this question. Its proof will make use of the following lemma.

LEMMA 1. *Let B be an $m \times m$ doubly stochastic matrix with row and column sums equal to T. Let A be a $u \times v$ submatrix of B and let B be partitioned into submatrices A, A_1, A_2, A_3 as in Figure 1. If S is the sum of all the elements in A and S_2 is the sum of all the elements in A_2, then*

$$S - S_2 = (u + v - m)T.$$

Proof. Let S_1 and S_3 be the sums of the elements in A_1 and A_3 respectively. We have

$$\begin{aligned} S + S_1 &= uT \\ S + S_3 &= vT \\ S + S_1 + S_2 + S_3 &= mT, \end{aligned}$$

from which the result follows.

THEOREM 2. *Let $A = (a_{ij})$ be an $n \times n$ matrix of non-negative real numbers and let M be the maximum row or column sum and S the sum of all the entries. If $r \leq n - 1$, the necessary and sufficient condition that there exist a matrix B which is an (r, T) d.s. extension of A is that $M \leq T \leq S/(n - r)$.*

Proof. If B is an (r, T) d.s. extension of A, we apply Lemma 1 to B. We have $S - S_2 = (n - r)T$. Since $0 \leq S_2$, it follows that

$$M \leq T = \frac{S - S_2}{n - r} \leq \frac{S}{n - r}.$$

Clearly, $T = S/(n - r)$ if, and only if, $S_2 = 0$ and $T = M$ if and only if $S_2 = S - (n - r)M$. To show the possibility of such extreme d.s. extensions we construct the appropriate matrices. We first construct a matrix $C = (c_{ij})$ which is an $(r, S/(n - r))$ d.s. extension of A. Let

$$c_{ij} = a_{ij} \qquad \text{for } i \leq n,\ j \leq n,$$

$$c_{ij} = \frac{\dfrac{S}{n - r} - R_i}{r} \qquad \text{for } i \leq n,\ n + 1 \leq j \leq n + r,$$

$$c_{ij} = \frac{\dfrac{S}{n - r} - C_j}{r} \qquad \text{for } n + 1 \leq i \leq n + r,\ j \leq n,$$

$$c_{ij} = 0 \qquad \text{for } n + 1 \leq i \leq n + r,\ n + 1 \leq j \leq n + r.$$

We next construct a matrix $D = (d_{ij})$ which is an (r, M) d.s. extension of A. Let

$$d_{ij} = a_{ij} \qquad \text{for } i \leq n,\ j \leq n,$$

$$d_{ij} = \frac{M - R_i}{r} \qquad \text{for } i \leq n,\ n + 1 \leq j \leq n + r,$$

$$d_{ij} = \frac{M - C_j}{r} \qquad \text{for } n + 1 \leq i \leq n + r,\ j \leq n,$$

$$d_{ij} = \frac{S - M(n - r)}{r^2} \qquad \text{for } n + 1 \leq i \leq n + r \text{ and } n + 1 \leq j \leq n + r.$$

For any T, $M \leq T \leq S/(n - r)$, let p be the unique real number $0 \leq p \leq 1$ defined by $pS/(n - r) + (1 - p)M = T$. The matrix $B = pC + (1 - p)D$ is an (r, T) d.s. extension of A.

We now define *stochastic rank*. An $n \times n$ matrix A with non-negative entries has stochastic rank σ if A can be embedded in a d.s. matrix B formed

from A by the addition of $n - \sigma$ rows and columns and if A cannot be embedded in a d.s. matrix B by the addition of fewer than $n - \sigma$ rows and columns. By Theorem 2, the least r for which A can be embedded in an $(n + r) \times (n + r)$ d.s. matrix B is the minimum r for which $M/S \leqslant 1/(n - r)$. This minimum r is $n - [S/M]$. It follows that $\sigma = [S/M]$.

An $n \times n$ *sub-permutation matrix of rank* r is an $n \times n$ matrix consisting of r ones, no two of which are in the same row or column, and $n^2 - r$ zeros. The convex hull of the sub-permutation matrices $P_k^{(r)}$ of rank r consists of all matrices A expressible in the form $A = \sum_k \lambda_k P_k^{(r)}$ where $\sum_k \lambda_k = 1$ and $\lambda_k \geqslant 0$ for all k. The convex polyhedral cone generated by the sub-permutation matrices $P_k^{(r)}$ of rank r consists of all matrices A expressible in the form $A = \sum_k \mu_k P_k^{(r)}$ where $\mu_k \geqslant 0$ for all k. In **(2)**, the authors showed that the necessary and sufficient condition that a matrix A of non-negative entries is in the convex hull of sub-permutation matrices of rank $n - r$ is that $S = n - r$ and $M \leqslant 1$. A simple restatement of this theorem is that the necessary and sufficient condition that a matrix of non-negative entries is in the convex polyhedral cone generated by the sub-permutation matrices of rank $n - r$ is that $M/S \leqslant 1/(n - r)$. Hence, the maximum rank $n - r$ satisfying this inequality is $[S/M]$ and this is equal to the stochastic rank σ of A. Thus, we have the following corollary to Theorem 2.

COROLLARY. *The stochastic rank of an $n \times n$ matrix A of non-negative entries is σ if, and only if, A is in the convex polyhedral cone of the $n \times m$ sub-permutation matrices of rank σ but is not in the convex polyhedral cone of the $n \times n$ sub-permutation matrices of rank $\sigma + 1$.*

4. Vertices of a set of doubly stochastic extensions. If we consider each (r, T) d.s. extension of a matrix A as a point in a space of dimension $(n + r)^2$, it is apparent the set α of all such matrices is convex. An *extreme or vertex matrix* for the convex set α is an (r, T) d.s. extension of A which is not expressible in the form $pC + (1 - p)D$ in which C and D belong to α, $C \neq D$ and $0 < p < 1$.

We may define the bipartite graph K_A of an $n \times m$ matrix A of non-negative entries to be the graph in which the vertex sets are the set of indices of the n rows and m columns and the edges are the places of the matrix in which the entries are positive. A graph is *disjoint* if no two of its edges have a vertex in common. A graph K_1 is a subgraph of K_2 if every edge of K_1 is an edge of K_2.

A *cycle* in a bipartite graph K is a finite subgraph K^1 with the following properties. Let I and J be the vertex sets. If (i_1, j_1) is any edge of K^1 then there exists exactly one vertex $i_2 \in I$, $i_2 \neq i_1$, such that (i_2, j_1) is an edge of K^1, and there exists exactly one vertex $j_2 \in J$, $j_2 \neq j_1$, such that (i_2, j_2) is an edge of K^1, and there exists exactly one vertex $i_3 \in I$, $i_3 \neq i_2$, such that (i_3, j_2) is an edge of K^1, etc. If after $2k - 1$ such steps, $k \geqslant 2$, we find that

(i_1, j_1), (i_2, j_1), (i_2, j_2), $\ldots$, (i_k, j_k), (i_1, j_k) are distinct and are exactly the edges of K^1, then K^1 is a cycle. It follows that for a cycle K^1 in the bipartite graph of a matrix, there exists no row or column which contains exactly one edge of the cycle.

In **(1)** the core of an R and C marking of an incidence matrix consists of the union of a number of cycles no two of which have an edge in common.

In Theorem 3 we require the following lemma.

LEMMA 2. *For a bipartite graph K, a necessary and sufficient condition that there exist a subgraph of K which is a cycle is that there exist a finite subgraph of K in which no vertex of either vertex set is edge connected to exactly one vertex of the other vertex set.*

Proof. The necessity is immediate.

To establish the sufficiency, we show that any finite subgraph K^1 of K in which no vertex of either vertex set is edge connected to exactly one vertex of the other, contains a subgraph which is a cycle of K. Let the vertex sets of K be I and J. If (i_1, j_1) is an edge of K^1, $i_1 \in I$ and $j_1 \in J$, there exists $i_2 \neq i_1$, such that (i_2, j_1) is an edge of K^1. Similarly, there exists $j_2 \neq j_1$ such that (i_2, j_2) is an edge of K^1. Continuing this process, since K^1 is a finite graph, it follows that in the sequence (i_1, j_1), (i_2, j_1), $(i_2, j_2,)\ldots$, there must exist a first edge E_1 in which either the i is identical with the i of a previous edge or the j is identical with the j of some previous edge. In either case, let this previous edge be E_0. The sequence of edges beginning with E_0 and ending with E_1 is a cycle.

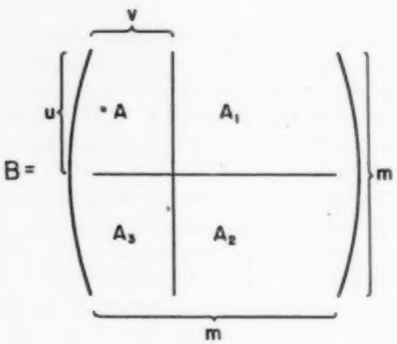

FIGURE 1

Now, consider any (r, T) d.s. extension B of a matrix A of non-negative elements and let B be partitioned into submatrices A, A_1, A_2, and A_3 as in Figure 1. Let

$$K_{A_1},\ K_{A_2},\ K_{A_3},$$

be the bipartite graphs of A_1, A_2, and A_3 and let L_B be the union of

$$K_{A_1},\ K_{A_2},\ K_{A_3},$$

so that K_B is the union of K_A and L_B. We are now in a position to state the main theorem of this section.

THEOREM 3. *Let α be the convex set of all (r, T) d.s. extensions of a matrix A. A necessary and sufficient condition that a matrix $B \in \alpha$ be a vertex matrix of the convex set α is that no subgraph of L_B is a cycle.*

Proof. If a subgraph $L_B{}^1$ of L_B is a cycle, let the edges of the cycle be $(i_1, j_1), (i_2, j_1), (i_2, j_2), \ldots, (i_k, j_k), (i_1, j_k)$.

Let $\epsilon = \frac{1}{2} \min b_{ij}$ taken over all edges (i, j) of $L_B{}^1$. Now, if $C = (c_{ij})$ is defined by

$$\begin{aligned} c_{ij} &= b_{ij} && \text{if } (i,j) \text{ is not an edge of } L_B{}^1, \\ c_{ij} &= b_{ij} + \epsilon && \text{if } (i,j) \text{ is } (i_1, j_1), (i_2, j_2), \ldots, \text{or } (i_k, j_k), \\ c_{ij} &= b_{ij} - \epsilon && \text{if } (i,j) \text{ is } (i_2, j_1), (i_3, j_2), \ldots, (i_1, j_k), \end{aligned}$$

and if $D = (d_{ij})$ is defined by

$$\begin{aligned} d_{ij} &= b_{ij} && \text{if } (i,j) \text{ is not an edge of } L_B{}^1 \\ &= b_{ij} - \epsilon && \text{if } (i,j) \text{ is } (i_1, j_1), (i_2, j_2), \ldots, \text{or } (i_k, j_k) \\ &= b_{ij} + \epsilon && \text{if } (i,j) \text{ is } (i_2, j_1), (i_3, j_2), \ldots, (i_1, j_k), \end{aligned}$$

clearly C and D belong to α. Since $B = \frac{1}{2}C + \frac{1}{2}D$, B is not a vertex matrix of the set α.

We now show that if B and C are (r, T) d.s. extensions of A such that $B \neq C$ and $K_B = K_c$ then L_B $(= L_C)$ has a subgraph $L_B{}^1$ which is a cycle. Indeed, let $L_B{}^*$ be the subgraph consisting of the edges (i, j) at which $0 < b_{ij}$, $0 < c_{ij}$ and $c_{ij} \neq b_{ij}$. Since $b_{ij} = c_{ij}$ for all (i, j) in K_A, $L_B{}^*$ is a subgraph of L_B and since $B \neq C$, $L_B{}^*$ has at least one edge. Since the matrices B and C are doubly stochastic with row and column sums equal to T, they cannot differ at exactly one place in a row or column. By Lemma 2, L_B (in fact $L_B{}^*$) contains a subgraph $L_B{}^1$ which is a cycle.

Next, suppose that $B \in \alpha$ is not a vertex, so that B is expressible in the form $B = pC + (1 - p)D$ where $0 < p < 1$, $C \neq D$ and C and D belong to α. Now L_C and L_D are subgraphs of L_B, but we cannot say $L_C = L_D = L_B$, for we might have $c_{ij} = 0$, $b_{ij} \neq 0$, and $d_{ij} \neq 0$. However, if $q \neq p$, $0 < q < 1$, then $E = qC + (1 - q)D$ belongs to α, $B \neq E$ and $K_B = K_E$. Hence L_B contains a subgraph $L_B{}^1$ which is a cycle. This completes the proof of Theorem 3.

COROLLARY. *Let α be the convex set of all (r, T) d.s. extensions of a matrix A. A necessary and sufficient condition that a matrix $B \in \alpha$ be a vertex matrix of the convex set α, is that there exist no matrix $C \in \alpha$ such that $B \neq C$ and $K_B = K_C$.*

LEMMA 3. *If P is an $r \times r$ matrix of non-negative elements with at least two non-zero elements in every row then the bipartite graph K_P contains a subgraph K_P^1 which is a cycle.*

Proof. Delete from P all the columns containing no non-zero elements and let the deleted $r \times s$ matrix $(s < r)$ be Q. If there are at least two non-zero elements in every column of Q then the required cycle exists by Lemma 2. If a column contains exactly one non-zero element q_{ij}, delete the ith row and jth column of Q and denote the deleted matrix by Q_1. Continue this process. If we find Q_t such that every column of Q_t contains 2 non-zero elements, the cycle exists by Lemma 2. If no such Q_t exists for $t = 1, 2, \ldots, s - 3$, then, Q_{s-2} is an $(r - s + 2) \times 2$ matrix with two columns and with two non-zero elements in every row. Since $r - s + 2 \geqslant 2$ the graph of Q_{t-2} contains a cycle.

Let B be an (r, T) d.s. extension of A. Let the rows and columns of B be rearranged as in Figure 1. If in the ith row of B $(i = 1, 2, \ldots, n)$ there is at most one $j > n$ such that the element $b_{ij} > 0$ then the ith row of A is *simply extended*. Similarly, if in the jth column of B there is at most one $i > n$ such that $b_{ij} > 0$, then the jth column of A is *simply* extended.

THEOREM 4. *If B is a vertex of the convex set α of all (r, T) d.s. extensions of A, then at least $n - r + 1$ of the rows and at least $n - r + 1$ of the columns of A are simply extended.*

Proof. Suppose that r rows of B are not simply extended. In each of these rows we have at least two elements

$$b_{ij_1} > 0 \quad \text{and} \quad b_{ij_2} > 0, \qquad j_1 > n, j_2 > n, j_1 \neq j_2.$$

Thus the $n \times r$ matrix A_1 (see Figure 1) contains an $r \times r$ submatrix A_1 in every row of which there are two non-zero elements.

Hence, by Lemma 3, the graph of

$$K_{A_1}$$

contains a subgraph which is a cycle and, by Theorem 3, B is not a vertex of α.

The proof when r columns of A are not simply extended is similar.

5. The connection between term rank and stochastic rank.

Since ρ is greater than or equal to ρ the least integer which is greater than or equal to S/M and since $\sigma = [S/M]$, we have the following result. If S/M is an integer, $\rho \geqslant \sigma$, and if S/M is not an integer, $\rho \geqslant \sigma + 1$.

For an $n \times n$ doubly stochastic matrix, $\rho = \sigma = n$, and for a sub-permutation matrix of rank r, $\rho = \sigma = r$. However, there are $n \times n$ matrices for which $\rho - \sigma = n - 1$. In fact, the matrix $A = (a_{ij})$ in which $a_{11} = n$, $a_{22} = a_{33} = \ldots = a_{nn} = 1$, $a_{ij} = 0$ for $i \neq j$ is such a matrix. We have $S/M = 2 - 1/n$. Thus $\sigma = 1$ and $\rho = n$.

For a matrix of zeros and ones, Ryser **(4;5)** has considered the transformation which replaces a minor

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{by} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The effect of this transformation is that the term rank varies between limits which Ryser finds. It is interesting to note that the stochastic rank of a matrix of zeros and ones is invariant under Ryser's transformation.

If $M < S/(n-r)$, then, since $\rho \geqslant \sigma \geqslant n - r$ there exist integers t such that $\rho \geqslant t \geqslant n - r$. We have the following theorem.

THEOREM 5. *Let A be an $n \times n$ matrix of non-negative elements. If $M < S/(n-r)$ and $M \leqslant T \leqslant S/(n-r)$ and if $K_A{}^1$ is any disjoint subgraph of K_A consisting of t edges ($\rho \geqslant t \geqslant n - r$) then there exists an (r, T) d.s. extension B of A with the property that the graph K_B contains a disjoint subgraph $K_B{}^1$ consisting of $n + r$ edges such that the edges common to $K_B{}^1$ and K_A are exactly those of $K_A{}^1$.*

Proof. If we select any p such that $0 < p < 1$, then since $M < S/(n-r)$, the matrix $B = pC + (1-p)D$ of Theorem 2 is an (r, T) d.s. extension of A in which every element of A_1, A_2, and A_3 (Figure 1) is positive. Thus all the places of A_1, A_2, and A_3 are edges of K_B. Rearrange the rows and columns of B so that $K_A{}^1$ consists of the places $(1, 1)$ $(2, 2)$ $(3, 3) \ldots (t, t)$. Now consider the disjoint graph L which has as its edges the places (i, j) of B defined by $i + j = n + t + r + 1$. Since $t \geqslant n - r$, we have $i + j \geqslant 2n + 1$ and hence every edge in L is a place in A_1, A_2, and A_3 and L is a subgraph of K_B. The number of edges in L is $n + r - t$. For an edge (i, j) of L we cannot have $i \leqslant t$, for this would imply $j \geqslant n + r + 1$ and similarly we cannot have $j \leqslant t$. Thus the edges of L and $K_A{}^1$ have no vertices in common. Clearly, the graph $K_B{}^1$ defined as the union of L and $K_A{}^1$ is the required disjoint subgraph of K_B.

Let K be a bipartite graph whose edges are a set of places in an $n \times n$ array and let A be a matrix formed by putting positive entries in the places of K and zeros elsewhere. For a given graph K, the term rank ρ of all such matrices A is the same and is equal to the exterior dimension **(3)** of the graph. Thus, term rank is really a graphical concept. On the other hand, for a given graph K, the stochastic rank σ of such matrices A will vary between 1 and an attainable maximum which we denote by σ_K. We now show that $\sigma_K \leqslant \rho \leqslant \sigma_K + 1$. The inequality on the left is a consequence of Theorems 1 and 2. To establish the inequality on the right, consider the matrix A formed by placing 1 in each of the ρ places of a maximal disjoint subgraph of K and ϵ in the other places of K. If a is the maximum number of places of K in any row or column of the $n \times n$ array and if b is the number of places in K, then

$$\frac{S}{M} = \frac{\rho + (b - \rho)\epsilon}{1 + (a-1)\epsilon}.$$

If $a = 1$, then $b = \rho$ and $\sigma_K = \rho$. In other cases, ϵ can be chosen small enough that $\sigma = [S/M] \geq \rho - 1$. Hence, $\sigma_K \geq \sigma \geq \rho - 1$ or $\rho \leq \sigma_K + 1$. The inequality $\sigma_K \leq \rho \leq \sigma_K + 1$ is best possible in the sense that there exist graphs K for which $\sigma_K = \rho$ and others for which $\rho = \sigma_K + 1$. The graph K consisting of the places on a main diagonal in an $n \times n$ array is a graph in which $\sigma_K = \rho$. The graph K consisting of 3 of the 4 places in a 2×2 array is such that $\rho = 2$. But any matrix A with non-zero elements in the places of K and a zero in the fourth place of the array lies in the convex polyhedral cone of sub-permutation matrices of rank 1 and does not lie in the convex polyhedral cone of sub-permutation matrices of rank 2. Hence $\sigma_K = 1$. Consider a graph K for which the maximum σ_K is attained in a matrix A in which S/m is non integral. We have $\rho \leq \sigma_K + 1$ and $\rho \geq \sigma_K + 1$ so that $\rho = \sigma_K + 1$.

The result just proved may be reformulated as the following theorem.

THEOREM 6. *Let K be a bipartite graph whose edges are the places in an $n \times n$ array. Let α be the set of all matrices A with positive entries in the places of K and zeros elsewhere. Let σ_K be the maximum stochastic rank attainable by a matrix of the set α. Then every matrix A of α has the same term rank ρ. Furthermore, if S_A and M_A represent the entry sum and maximal row or column sum of A respectively then*

$$\rho = \sup_{A \epsilon \alpha} \left(\frac{S_A}{M_A}\right).$$

Also if this supremum is attained by some matrix A, then $\rho = \sigma_K$, otherwise $\rho = \sigma_K + 1$.

6. Linear programming formulation. Some of the theorems concerning (r, T) d.s. extensions of an $n \times n$ matrix A may be reformulated as problems in the language of linear programming. In these reformulations the restrictions on A to non-negative entries may be relaxed somewhat. The only requirement is that A satisfy the condition $S \geq (n - r)M \geq 0$. Two such formulations follow.

PROBLEM 1. *Let A be an $n \times n$ matrix having $S \geq (n - r)M \geq 0$, and let T be any number. Find a set of numbers χ_{ij} $(i = 1, 2, \ldots, n + r; j = 1, 2, \ldots, n + r$; at least one of i and j is greater than n), subject to the following conditions.*

(1) $\chi_{ij} \geq 0$ for all i, j.

(2) $\sum_{j=1}^{n} a_{ij} + \sum_{j=n+1}^{n+r} \chi_{ij} = T$ for $i = 1, 2, \ldots, n$.

(3) $\sum_{j=1}^{n+r} \chi_{ij} = T$ for $i = n + 1, n + 2, \ldots, n + r$.

(4) $\sum_{i=1}^{n} a_{ij} + \sum_{i=n+1}^{n+r} \chi_{ij} = T$ for $j = 1, 2, \ldots, n$.

(5) $\sum_{i=1}^{n+r} \chi_{ij} = T$ for $j = n + 1, n + 2, \ldots, n + r$.

Theorem 2 states that the inequalities have solutions if and only if $M \leqslant T \leqslant S/(n-r)$ and exhibits some of these solutions. If now each set of values of χ_{ij} satisfying (1), (2), . . . , (5) is considered as a point in a space of $(n+r)^2 - n^2$ dimensions, the set of all such points is convex and Theorem 3 gives a graphical characterization of the vertices of this set.

PROBLEM 2. *Let A be an $n \times n$ matrix having $S \geqslant (n-r)M \geqslant 0$. Find a set of numbers χ_{ij} ($i = 1, 2, \ldots, n+r$; $j = 1, 2, \ldots, n+r$; at least one of i and j is greater than n), subject to the following conditions:*

(1) $\chi_{ij} \geqslant 0$ for all i, j.

(2) $\sum_{j=1}^{n} a_{ij} + \sum_{j=n+1}^{n+r} \chi_{ij} = \sum_{j=1}^{n+r} \chi_{n+r,j}$ for $i = 1, 2, \ldots, n$.

(3) $\sum_{j=1}^{n+r} \chi_{ij} = \sum_{j=1}^{n+r} \chi_{n+r,j}$ for $i = n+1, n+2, \ldots, n+r-1$.

(4) $\sum_{i=1}^{n} a_{ij} + \sum_{i=n+1}^{n+r} \chi_{ij} = \sum_{i=1}^{n+r} \chi_{i,n+r}$ for $j = 1, 2, \ldots, n$.

(5) $\sum_{i=1}^{n+r} \chi_{ij} = \sum_{i=1}^{n+r} \chi_{i,n+r}$ for $j = n+1, n+2, \ldots, n+r-1$.

The sum

$$\sum_{i=1}^{n+r} \chi_{i,n+r}$$

is to be maximized or minimized.

In this formulation our theories state that feasible solutions always exist for both the maximum and minimum problems. They also exhibit solutions at which the maximum and minimum are attained and state that the maximum value is $S/(n-r)$ and the minimum value is M. Our graphical theorems characterize the sets of all maximal and minimal solutions.

REFERENCES

1. A. L. Dulmage and N. S. Mendelsohn, *Some generalizations of the problem of distinct representatives*, Can. J. Math., *10* (1958), 230–41.
2. ——— *The convex hull of sub-permutation matrices*, Proc. Amer. Math. Soc., *9* (1958), 253–4.
3. ——— *Coverings of bipartite graphs*, Can. J. Math., *10* (1958), 517–34.
4. H. J. Ryser, *Combinatorial properties of matrices of zeros and ones*, Can. J. Math., *9* (1957), 371–7.
5. ——— *The term rank of a matrix*, Can. J. Math., *10* (1957), 57–65.

University of Manitoba

DISJOINT TRANSVERSALS OF SUBSETS

P. J. HIGGINS

1. Introduction. Let $A_1, A_2, \ldots, A_n$ be a finite collection of subsets (not necessarily distinct) of a set A. By a *transversal*[1] of $A_1 A_2, \ldots, A_n$ we shall mean a set of n distinct elements $a_1, a_2, \ldots, a_n$ of A such that, for some permutation $i_1, i_2, \ldots, i_n$ of the integers $1, 2, \ldots, n$,

$$a_j \in A_{i_j} \qquad (j = 1, 2, \ldots, n).$$

More generally, we shall say that the set $\{a_1, a_2, \ldots, a_r\}$, $(r < n)$ is a *partial transversal* of $A_1, A_2, \ldots, A_n$ of *length* r if (i) $a_1, a_2, \ldots, a_r$ are distinct elements of A and (ii) there exists a set of distinct integers $i_1, i_2, \ldots, i_r$ such that

$$a_j \in A_{i_j} \qquad (j = 1, 2, \ldots, r).$$

A well-known theorem of P. Hall **(2)** states that the sets $A_1, A_2, \ldots, A_n$ have a transversal (of length n) if, and only if, every k of them contain collectively at least k distinct elements $(k = 1, 2, \ldots, n)$. A generalization of this theorem by Ore **(3)** states that the sets $A_1, A_2, \ldots, A_n$ have a partial transversal of length $r < n$ if, and only if, every k of them contain collectively at least $k + r - n$ distinct elements $(n - r + 1 < k < n)$.

In this paper we enquire under what conditions the sets $A_1, A_2, \ldots, A_n$ will have m mutually disjoint partial transversals of prescribed lengths $r_1, r_2, \ldots, r_m$. As in the two theorems quoted above, the obvious necessary conditions are again found to be sufficient. As a special case we deduce a theorem of Ryser **(4)** and Gale **(1)** concerning the existence of matrices of 0's and 1's with prescribed row sums and column sums.

2. Notation. Throughout our argument n will denote a fixed positive integer (the number of subsets A_j), and $r_1, r_2, \ldots, r_m$ will denote positive integers not exceeding n. We shall suppose that $r_1 > r_2 > \ldots > r_m > 0$ and think of these integers as a partition $[r_i]$ of $r_1 + r_1 + \ldots + r_m$. The conjugate partition $[r_j^*]$ is defined as usual:

$$(1) \qquad r_j^* = \sum_{r_i > j} 1 \qquad (j = 1, 2, \ldots, r_1).$$

It is convenient also to define $r_j^* = 0$ if $r_1 < j < n$, which is in accord with (1) if we interpret empty sums as zero.

Received May 22, 1958.

[1]This term, due to P. Hall, is normally used when the sets $A_1, A_2, \ldots A_n$ are disjoint, but its use in this wider sense is convenient here.

We now write

$$\alpha_k = \sum_{j=n-k+1}^{n} r_j^* \qquad (k = 1, 2, \ldots, n). \tag{2}$$

An alternative expression for α_k can be obtained as follows. If s and t are integers, let $Ex(s, t)$ denote the excess, if any, of s over t, that is, $Ex(s, t) = s - t$ if $s > t$, and $Ex(s, t) = 0$ if $s < t$. Then

$$\alpha_k = \sum_{i=1}^{m} Ex(r_i, n - k) \qquad (k = 1, 2, \ldots, n). \tag{3}$$

The easiest way to see this is to draw a partition diagram for $[r_i]$, that is, an $m \times n$ matrix whose ith row has entries 1 in the first r_i places and 0 elsewhere. Then r_j^* is the number of 1's in the jth column, and α_k is the number of 1's in the last k columns. However, the ith row has exactly $Ex(r_i, n - k)$ 1's in the last k columns, and (3) follows.

3. Disjoint partial transversals. Suppose that $A_1, A_2, \ldots, A_n$ have disjoint partial transversals (D.P.T.'s) $R_1, R_2, \ldots, R_m$ of lengths $r_1, r_2, \ldots, r_m$ respectively. The elements of R_i represent r_i of the A's. Of these A's at least $Ex(r_i, n - k)$ must be included in any collection of k of the A's. It follows that every k of the A's must contain between them at least $Ex(r_i, n - k)$ distinct elements out of R_i and hence at least

$$\alpha_k = \sum_{i=1}^{m} Ex(r_i, n - k)$$

distinct elements altogether, since the R's are disjoint. Our theorem asserts that this necessary condition is also sufficient.

THEOREM. *A necessary and sufficient condition for $A_1, A_2, \ldots, A_n$ to have mutually disjoint partial transversals of lengths $r_1, r_2, \ldots, r_m$ is that, for $k = 1, 2, \ldots, n$, every k of the A's contain between them at least α_k distinct elements, where α_k is defined by* (1) *and* (2) *above.*

We observe here that the case $m = 1, r_1 = r$, is precisely Ore's theorem since we then have $\alpha_k = 0$ $(1 < k < n - r)$ and $\alpha_k = k + r - n$ $(n - r + 1 < k < n)$.

The proof of sufficiency proceeds by induction on n. It is trivial when $n = 1$, and from now on we shall assume the result for all collections of $n' < n$ sets and all sets of integers $r_i < n'$.

We distinguishtwo cases which are mutually exclusive and cover all possibilities:

Case 1. $m > 2$ and $r_m < n$, $r_{m-1} < n$;

Case 2. $r_1 = r_2 = \ldots = r_{m-1} = n$, $1 < r_m < n$.

We shall first reduce Case 1 to Case 2.

If $[r_i]$ is a partition falling under Case 1, we define a new partition $[\bar{r}_i]$, the *reduction* of $[r_i]$, as follows. Let $r_1 = r_2 = \ldots = r_t = n$, $r_{t+1} < n$, where, by assumption, $0 \leqslant t \leqslant m - 2$. Then $\bar{r}_m = r_m - 1$, $\bar{r}_{t+1} = r_{t+1} + 1$, and $\bar{r}_i = r_i$ for all other values of i. Clearly $\bar{r}_1 \geqslant \bar{r}_2 \geqslant \ldots \geqslant \bar{r}_m$, and by a finite number of such reductions any partition falling under Case 1 can be reduced to one falling under Case 2. Note that we may have $\bar{r}_m = 0$, in which case the value of m is reduced by 1. It will, however, be convenient at times to retain such vanishing parts of a partition and interpret a partial transversal of length zero as the empty set. This will not affect the proof in any way.

To reduce Case 1 to Case 2 it is enough to prove

LEMMA. *If the theorem is true for the partition $[\bar{r}_i]$ then it is also true for the partition $[r_i]$.*

Proof. Suppose that $[r_i]$ falls under Case 1, and every k of the A's contain between them at least α_k distinct elements $(k = 1, 2, \ldots, n)$.

Case 1 (a). First consider the possibility that for some k $(1 \leqslant k \leqslant n - 1)$ there is a collection of k of the A's, say $A_1, A_2, \ldots, A_k$, which contain between them precisely α_k distinct elements. We construct two new partitions $[p_i]$ and $[q_i]$ where $p_i = Ex(r_i, n - k)$, $q_i = \min(r_i, n - k)$ $(i = 1, 2, \ldots, m)$. Then $p_i + q_i = r_i$ $(i = 1, 2, \ldots, m)$, $p_j^* = r_{j+n-k}^*$ $(j = 1, 2, \ldots, k)$, and $q_j^* = r_j^*$ $(j = 1, 2, \ldots, n - k)$. We apply our induction hypothesis to the sets $A, A_2, \ldots, A_k$ with the partition $[p_i]$ and to the sets $A_{k+1}, A_{k+2}, \ldots, A_n$ with the partition $[q_i]$. For this purpose let β_s and γ_s be the integers obtained from $[p_i]$ and $[q_i]$ in the same way that the α_s were obtained from $[r_i]$. Thus

$$\beta_s = \sum_{j=k-s+1}^{k} p_j^* = \sum_{j=n-s+1}^{n} r_j^* = \alpha_s \qquad (s = 1, 2, \ldots, k),$$

and

$$\gamma_s = \sum_{j=n-k-s+1}^{n-k} q_j^* = \sum_{j=n-k-s+1}^{n-k} r_j^* = \alpha_{k+s} - \alpha_k \qquad (s = 1, 2, \ldots, n - k).$$

By assumption, every s of the sets $A_1, A_2, \ldots, A_k$ contain between them at least $\alpha_s = \beta_s$ distinct elements. Also, any s of the sets $A_{k+1}, A_{k+2}, \ldots, A_n$ contain, together with all of $A_1, A_2, \ldots, A_k$, at least α_{k+s} distinct elements. Since $A_1 \cup A_2 \cup \ldots \cup A_k$ contains precisely α_k elements, any s of the sets $A_{k+1}, A_{k+2}, \ldots, A_n$ must contain between them at least $\alpha_{k+s} - \alpha_k = \gamma_s$ distinct elements not in $A_1 \cup A_2 \cup \ldots \cup A_k$. It follows that there exist D.P.T.'s $P_1, P_2, \ldots, P_m$ of $A_1, A_2, \ldots, A_k$ of lengths $p_1, p_2, \ldots, p_m$, and D.P.T.'s $Q_1, Q_2, \ldots, Q_m$ of $A_{k+1}, A_{k+2}, \ldots, A_n$ of lengths $q_1, q_2, \ldots, q_m$, none of the Q's having any elements in common with any of the P's. The sets $P_1 \cup Q_1, P_2 \cup Q_2, \ldots, P_m \cup Q_m$ are then D.P.T.'s of $A_1, A_2, \ldots, A_n$ of lengths $r_1, r_2, \ldots, r_m$.

Case 1 (b). If no such collection of A's exists then, for $k = 1, 2, \ldots, n - 1$, every k of the A's must contain between them at least $\alpha_k + 1$ distinct elements,

and we now appeal to the reduced partition. We observe that in passing from $[r_i]$ to $[\bar{r}_i]$ one of the r_j^* is increased by 1, and one of them is decreased by 1, the others being unaltered. Hence, in the obvious notation,

$$\bar{\alpha}_k = \sum_{j=n-k+1}^{n} \bar{r}_j^* \leqslant 1 + \sum_{j=n-k+1}^{n} r_j^* = 1 + \alpha_k \qquad (k = 1, 2, \ldots, n),$$

while $\bar{\alpha}_n = \alpha_n$. Thus every k of the A's contain between them at least $\bar{\alpha}_k$ distinct elements $(k = 1, 2, \ldots, n)$. Assuming the theorem for the partition $[\bar{r}_i]$, we can find D.P.T.'s $\bar{R}_1, \bar{R}_2, \ldots, \bar{R}_m$ of lengths $\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_m$. Now $\bar{r}_{t+1} > r_{t+1} \geqslant r_m > \bar{r}_m$ (t has the same meaning as before). Hence there must be in $\bar{R}_{t+1}$ at least one element which represents a set A_j not represented by any element of $\bar{R}_m$. If we transfer this element from $\bar{R}_{t+1}$ to $\bar{R}_m$ we obtain D.P.T.'s of lengths $r_1, r_2, \ldots, r_m$. This proves the lemma.

It remains to prove the theorem in Case 2, that is, under the assumptions $r_1 = r_2 = \ldots = r_{m-1} = n$, $1 \leqslant r_m < n$, $m > 1$. Then $r_j^* = m$ for $j = 1, 2, \ldots, r$, and $r_j^* = m - 1$ for $j = r + 1, r + 2, \ldots, n$, where for convenience we write $r_m = r$. We now make the further definition

$$\delta_k = \sum_{j=1}^{k} r_j^* \geqslant \alpha_k \qquad (k = 1, 2, \ldots, n).$$

Assume that $A_1, A_2, \ldots, A_n$ satisfy the conditions of the theorem.

Case 2 (a). First suppose that, for $k = 1, 2, \ldots, n - 1$, every k of the A's contain between them at least δ_k distinct elements. The same will be true for $k = n$ since $\delta_n = \alpha_n$. Consider a collection of sets $\{B_j\}$ consisting of m repetitions of each of the sets $A_1, A_2, \ldots, A_r$ and $m - 1$ repetitions of each of the sets $A_{r+1}, A_{r+2}, \ldots, A_n$ (if any). There are α_n sets altogether, and we shall show that, for $s = 1, 2, \ldots, \alpha_n$, any s of these sets contain between them at least s distinct elements. We must first count the number k of distinct* A's included amongst s of the B's. Clearly $k \geqslant s/m$; and if $s > mr$ then $k \geqslant r + (s - mr)/(m - 1)$. If $s \leqslant mr$, then $s/m \leqslant r$ and, if k' is the smallest integer such that $k' \geqslant s/m$, then $k' \leqslant r$. Hence $\delta_{k'} = k'm \geqslant s$, and any s of the B's must contain between them at least $\delta_k \geqslant \delta_{k'} \geqslant s$ distinct elements. On the other hand, if $s > mr$, then $k - r \geqslant (s - mr)/(m - 1)$, and $\delta_k = rm + (k - r)(m - 1) \geqslant rm + (s - mr) = s$. Thus again any s of the B's must contain between them at least s distinct elements. Applying Hall's theorem quoted in the introduction, we can find a complete transversal of the B's. The α_n distinct elements in this transversal comprise m distinct representatives of each of the sets $A_1, A_2, \ldots, A_r$ and $m - 1$ distinct representatives of each of the sets $A_{r+1}, A_{r+2}, \ldots, A_n$. It is easy to see that these elements can be arranged to form D.P.T.'s of $A_1, A_2, \ldots, A_n$, $m - 1$ of length n and one of length r.

*By "distinct" we mean here "having distinct suffixes." Thus distinct A's may have the same members.

Case 2 (b). The alternative to 2 (a) is that for some k $(1 \leqslant k \leqslant n-1)$ there is a collection of k of the A's whose union contains fewer than δ_k distinct elements (but at least α_k). From all such collections (for all possible values of k) we pick one collection consisting of, say, k A's whose union contains $\alpha_k + u$ distinct elements with *u as small as possible.* Thus every s of the A's $(s = 1, 2, \ldots, n)$ contain between them at least $\min(\delta_s, \alpha_s + u)$ distinct elements. (This statement for $s = n$ follows from the fact that $\delta_n = \alpha_n$.) Let the chosen sets be $A_1, A_2, \ldots, A_k$ (k is now fixed, $1 \leqslant k \leqslant n-1$). If $u = 0$ we may, of course, proceed as in Case 1 (a). This fails, however, if $u > 0$, and we must appeal again to the special form of the partition $[r_i]$.

Consider the sums of k successive r^*'s, that is, the integers $\epsilon_i = r_{i+1}^* + r_{i+2}^* + \ldots + r_{i+k}^*$ $(i = 0, 1, \ldots, n-k)$. Clearly $\delta_k = \epsilon_0 \geqslant \epsilon_1 \geqslant \ldots \geqslant \epsilon_{n-k} = \alpha_k$. Also $\epsilon_i - \epsilon_{i+1} \leqslant 1$ since $m = r_1^* \geqslant r_2^* \geqslant \ldots \geqslant r_n^* \geqslant m-1$. Now $\delta_k > \alpha_k + u > \alpha_k$; hence there is an integer t $(1 \leqslant t < n-k)$ such that $\epsilon_t = \alpha_k + u$. We may take $t \leqslant r$ since, if ϵ_r is defined, its value is $(m-1)k$ which must also be the value of α_k.

In the partition diagram of $[r_i]$ we now look at columns $t+1, t+2, \ldots, t+k$. They form the partition diagram of $[p_i]$ where $p_1 = p_2 = \ldots = p_{m-1} = k$, $p_m = r - t$, and

$$\sum_{i=1}^{m} p_i = \alpha_k + u.$$

The remaining columns form the partition diagram of $[q_i]$ where $q_1 = q_2 = \ldots = q_{m-1} = n-k$, $q_m = t$. The integers β_s and γ_s obtained from $[p_i]$ and $[q_i]$ in the same way that the α_s were obtained from $[r_i]$ are given by

$$\beta_s = \sum_{j=k-s+1}^{k} p_j^* = \sum_{j=t+k-s+1}^{t+k} r_j^* \qquad (s = 1, 2, \ldots, k),$$

$$\gamma_s = \sum_{j=n-k-s+1}^{n-k} q_j^* = \begin{cases} \alpha_s & \text{if } s \leqslant n-k-t \\ \alpha_{k+s} - (\alpha_k + u) & \text{if } n-k-t < s \leqslant n-k. \end{cases}$$

Consider a collection of $s \leqslant k$ of the sets $A_1, A_2, \ldots, A_k$. Between them they contain at least $\min(\delta_s, \alpha_s + u)$ distinct elements. Now

$$\delta_s = \sum_{j=1}^{s} r_j^* \geqslant \sum_{j=t+k-s+1}^{t+k} r_j^* = \beta_s.$$

Also

$$\begin{aligned} \alpha_s + u = (\alpha_k + u) - (\alpha_k - \alpha_s) &= \sum_{j=t+1}^{t+k} r_j^* - \sum_{j=n-k+1}^{n-s} r_j^* \\ &\geqslant \sum_{j=t+1}^{t+k} r_j^* - \sum_{j=t+1}^{t+k-s} r_j^* \qquad (\text{since } t \leqslant n-k) \\ &= \beta_s. \end{aligned}$$

Thus any s of $A_1, A_2, \ldots, A_k$ contain between them at least β_s distinct elements, and since $k < n$, we may apply our induction hypothesis to find D.P.T.'s $P_1, P_2, \ldots, P_m$ of $A_1, A_2, \ldots, A_k$ of lengths $p_1, p_2, \ldots, p_m$.

Now consider a collection of $s \leq n - k$ of the sets $A_{k+1}, A_{k+2}, \ldots, A_n$. Together with all of $A_1, A_2, \ldots, A_k$ they contain at least $\min(\delta_{k+s}, \alpha_{k+s} + u)$ distinct elements. Since $A_1 \cup A_2 \cup \ldots \cup A_k$ contains exactly $\alpha_k + u$ elements, the s sets from $A_{k+1}, A_{k+2}, \ldots, A_n$ must contain between them at least $\min(\delta_{k+s} - (\alpha_k + u), \alpha_{k+s} - \alpha_k)$ distinct elements not already used in $P_1, P_2, \ldots, P_m$. If we can show that, for $s = 1, 2, \ldots, n - k$,

$$\text{(i)} \quad \delta_{k+s} - (\alpha_k + u) \geq \gamma_s,$$

and

$$\text{(ii)} \quad \alpha_{k+s} - \alpha_k \geq \gamma_s,$$

we may apply our induction hypothesis to obtain D.P.T.'s $Q_1, Q_2, \ldots, Q_m$ of $A_{k+1}, A_{k+2}, \ldots, A_n$ of lengths $q_1, q_2, \ldots, q_m$ from elements not already used in $P_1, P_2, \ldots, P_m$. If $s > n - k - t$, these inequalities are obvious; for then $\gamma_s = \alpha_{k+s} - (\alpha_k + u)$, and clearly $\delta_{k+s} \geq \alpha_{k+s}$, $\alpha_k \leq \alpha_k + u$. On the other hand, if $s \leq n - k - t$, then $\gamma_s = \alpha_s$. In this case we observe that $\delta_{k+s} \geq \delta_k + \alpha_s > (\alpha_k + u) + \alpha_s$, from which (i) follows. Also $\alpha_{k+s} \geq \alpha_k + \alpha_s$, from which (ii) follows. This establishes the existence of $Q_1, Q_2, \ldots, Q_m$.

Finally, $P_1 \cup Q_1$, $P_2 \cup Q_2, \ldots, P_m \cup Q_m$ are D.P.T.'s of $A_1, A_2, \ldots, A_n$ of lengths $r_1, r_2, \ldots, r_m$, and the theorem is proved.

The application to matrices of 0's and 1's, mentioned in the introduction, is immediate. Let $n \geq r_1 \geq r_2 \geq \ldots \geq r_m \geq 0$ and $s_1 \geq s_2 \geq \ldots \geq s_n \geq 0$. The insertion of 1's in an $m \times n$ matrix so that there are at least r_i 1's in the ith row ($i = 1, 2, \ldots, m$) and not more than s_j in the jth column ($j = 1, 2, \ldots, n$) is equivalent to the construction of D.P.T.'s of lengths $r_1, r_2, \ldots, r_m$ of n *disjoint* sets containing respectively $s_1, s_2, \ldots, s_n$ elements. Our theorem gives as necessary and sufficient conditions for the existence of such D.P.T.'s the inequalities

$$\sum_{j=n-k+1}^{n} s_j \geq \alpha_k = \sum_{j=n-k+1}^{n} r_j^* \qquad (k = 1, 2, \ldots, n).$$

(The inclusion of zeros amongst the r's affects neither the hypotheses nor the conclusion.) If we require exactly r_i 1's in the ith row and exactly s_j in the jth column, we need only add the condition

$$\sum_{j=1}^{n} s_j = \sum_{i=1}^{m} r_i.$$

These are the conditions found by Ryser and Gale.

References

1. David Gale, *A theorem on flows in networks*, Pacific J. Math., 7 (1957), 1073–82.
2. Philip Hall, *On representatives of subsets*, J. Lond. Math. Soc., *10* (1935), 26–30.
3. Oystein Ore, *Graphs and matching theorems*, Duke Math. J., *22* (1955), 625–39.
4. H. J. Ryser, *Combinatorial properties of matrices of zeros and ones*, Can. J. Math., *9* (1957), 371–7.

Harvard University

SEPARATION AND APPROXIMATION IN TOPOLOGICAL VECTOR LATTICES

SOLOMON LEADER

1. Introduction. Spectral theory in its lattice-theoretic setting proves abstractly that the indicators of measurable sets generate the space L of Lebesgue-integrable functions on an interval. We are concerned here with abstractions suggested by the fact that indicators of intervals suffice to generate L. Our results show that the approximation of arbitrary elements of a topological vector lattice rests upon the ability to separate disjoint elements f and g by an operation that behaves in the limit like a projection annihilating f and leaving g invariant.

The introduction of this concept of separation together with the notion of limit unit leads (via the Fundamental Lemma) to abstract generalizations of the Radon-Nikodym Theorem (Theorem 1) and the Stone-Weierstrass Theorem (Theorem 3). Even for lattices which have representations as function spaces our abstract approach has several advantages: (i) the domain plays no explicit role in the theory, (ii) we are not restricted to the topology of uniform convergence, and (iii) the functions under consideration need not be bounded, although they must be limits of bounded functions. Thus, Theorem 3 is actually stronger than Stone's theorem **(12)**. We do not assume conditional σ-completeness **(1)** in our lattices, so countable-additivity plays no role in the Boolean ring of Theorem 1.

The author is indebted to the referees for clarifying the general setting of the theory.

2. Positive operators on a vector lattice. Let $\mathfrak{L}$ be a vector lattice with real scalars. The following lattice-group properties will prove useful **(1, 4, 9)**:

$$f + g = f \vee g + f \wedge g \tag{2.1}$$

$$(f - f \wedge g) \wedge (g - f \wedge g) = 0 \tag{2.2}$$

$$|f \wedge h - g \wedge h| \leq |f - g| \tag{2.3}$$

$$|f \vee h - g \vee h| \leq |f - g|. \tag{2.4}$$

An *operator* on $\mathfrak{L}$ is a linear mapping of $\mathfrak{L}$ into itself. The operators on $\mathfrak{L}$ are partially ordered by defining $P \leq Q$ whenever $Pf \leq Qf$ for all $f \geq 0$ in $\mathfrak{L}$. Thus, positive operators are order-preserving:

$$\text{If } P \geq 0 \text{ and } f \leq g, \text{ then } Pf \leq Pg. \tag{2.5}$$

Received June 5, 1958. The author is grateful for the support of the Research Council of Rutgers University.

A *contractor* is an operator P such that

(2.6) $$0 \leqslant P \leqslant I$$

where I is the identity operator. We shall use the abbreviation P' for $I - P$. Thus, P is a contractor if, and only if, both P and P' are positive operators. Note that P' is a contractor whenever P is a contractor, and PQ is a contractor whenever P and Q are contractors.

Contractors interest us because they commute with the lattice operations:

(2.7) $$P(f \wedge g) = Pf \wedge Pg$$

(2.8) $$P(f \vee g) = Pf \vee Pg$$

and

(2.9) $$P|f| = |Pf|.$$

To prove (2.7) let $h = Pf \wedge Pg$. Since $f \wedge g \leqslant f$ and $f \wedge g \leqslant g$, (2.5) gives $P(f \wedge g) \leqslant Pf$ and $P(f \wedge g) \leqslant Pg$. Hence $P(f \wedge g) \leqslant h$. To reverse this inequality we have $h \leqslant Pf$ and $P'(f \wedge g) \leqslant P'f$. Adding these gives $h + P'(f \wedge g) \leqslant f$. Similarly, $h + P'(f \wedge g) \leqslant g$. Hence $h + P'(f \wedge g) \leqslant f \wedge g$. Transposing the second term on the left gives $h \leqslant P(f \wedge g)$. Hence (2.7). The dual statement (2.8) follows from (2.7) and the identity (2.1). To obtain (2.9) set $g = -f$ in (2.8).

We call an idempotent contractor a *projector*. If A and B are projectors and $f \geqslant 0$, then

(2.10) $$ABf = Af \wedge Bf.$$

To derive (2.10) let $g = Af \wedge Bf$. Now $ABf \leqslant Bf \leqslant f$ by (2.6). Applying A to the latter inequality gives $ABf \leqslant Af$. Hence $ABf \leqslant g$. To reverse this inequality note that $0 \leqslant g \leqslant Af$ and $0 \leqslant g \leqslant Bf$. Since $A^2 = A$, $A'A = 0$, so $A'g = 0$ by (2.5). Thus $Ag = g$ and similarly $Bg = g$. Hence $ABg = Ag = g$. Since $Af \leqslant f$ and $Bf \leqslant f$, $g \leqslant f$. So $ABg \leqslant ABf$. That is, $g \leqslant ABf$. Hence (2.10).

From (2.10) it follows that projectors commute: $AB = BA$. Moreover, in terms of the operator ordering, (2.10) gives $A \cap B = AB$ and hence $A \cup B = A + B - AB$, which are easily seen to be projectors. Thus, the projectors on $\mathfrak{L}$ form a Boolean algebra with I as unit.

We remark that if $\mathfrak{L}$ is non-Archimedean, contractors need not commute.

3. Topological vector lattices. L is a *topological vector lattice* if it is a vector lattice with a topology making it a topological vector space possessing a local base of neighbourhoods $\mathfrak{N}$ of 0 such that

(3.1) $$f \text{ is in } \mathfrak{N} \text{ whenever } |f| \leqslant |g| \text{ for some } g \text{ in } \mathfrak{N}.$$

(In **(10)** $\mathfrak{L}$ is called a locally-solid lattice-ordered linear topological space.) The lattice operations as well as the vector operations are continuous in $\mathfrak{L}$. Every Banach lattice **(1)** is clearly a topological vector lattice.

Given an arbitrary set $\mathfrak{U}$ of elements in a topological vector space $\mathfrak{B}$, we say $\mathfrak{U}$ *generates* $\mathfrak{W}$ if $\mathfrak{W}$ is the smallest closed linear subspace of $\mathfrak{B}$ which contains $\mathfrak{U}$.

A positive element u in a topological vector lattice is a *limit bound* of f if

$$|f| \wedge nu \to |f| \qquad \text{as } n \to \infty. \tag{3.2}$$

f is *bounded* relative to u if $|f| \leqslant nu$ for some n. Now, u is a limit bound of f if, and only if, f is a limit of elements bounded relative to u. For, given (3.2) and $|h| \leqslant |f|$ we have, using (2.3), $0 \leqslant |f| \wedge nu - h \wedge nu \leqslant |f| - h$. Hence $0 \leqslant h - h \wedge nu \leqslant |f| - |f| \wedge nu$. From (3.2) and (3.1) we have $h \wedge nu \to h$. Taking first $h = f^+$ and then $h = f^-$ gives $f^+ \wedge nu - f^- \wedge nu \to f$ as $n \to \infty$. Conversely, given a net **(8)** of bounded elements converging to f, $f_\iota \to f$, we have $|f_\iota| = |f_\iota| \wedge nu$ for n sufficiently large. So using (2.3),

$$0 \leqslant |f| - |f| \wedge nu \leqslant \big||f| - |f_\iota|\big| + \big||f_\iota| - |f| \wedge nu\big| \leqslant 2\big||f| - |f_\iota|\big| \leqslant 2|f - f_\iota|.$$

Hence (3.2) follows from (3.1).

We say u is a *limit unit* in $\mathfrak{L}$ if u is a limit bound for every f in $\mathfrak{L}$, that is, if the bounded elements relative to u are dense in $\mathfrak{L}$. A limit unit is always a weak unit **(1)** if the topology in $\mathfrak{L}$ is T_1, that is, if finite sets are closed. To prove this let $f \wedge u = 0$. Then we have $1/n\ (f \wedge nu) \leqslant f$ and $1/n\ (f \wedge nu) \leqslant u$. So $f \wedge nu = 0$. Hence (3.2) implies $f = 0$. We remark that a weak unit need not be a limit unit.

A set $\mathfrak{C}$ of operators on a topological vector lattice $\mathfrak{L}$ is said to *separate f from g* if for every neighbourhood $\mathfrak{N}$ of 0 in $\mathfrak{L}$ there exists P in $\mathfrak{C}$ such that both $f - Pf$ and Pg are in $\mathfrak{N}$, that is, if there exists a net P_ι in $\mathfrak{C}$ such that $P_\iota f \to f$ and $P_\iota g \to 0$. We say $\mathfrak{C}$ *separates f and g* if it separates f from g and g from f.

4. Approximation by contractors on a limit unit.

Our approximation theorems all depend upon the following lemma:

FUNDAMENTAL LEMMA. *Let u be a limit unit in a topological vector lattice $\mathfrak{L}$ and $\mathfrak{C}$ a set of contractors on $\mathfrak{L}$ such that $\mathfrak{C}$ separates every pair f and g in $\mathfrak{L}$ for which $f \wedge g = 0$. Then the set of all $PQ'u$ with P and Q in $\mathfrak{C}$ generates $\mathfrak{L}$.*

Proof. Since u is a limit unit we need only show that for $|f| \leqslant \lambda u$ and $\mathfrak{N}$ any neighbourhood of 0 satisfying (3.1) there exists g of the form $\sum \lambda_k P_k Q_k' u$ with P_k and Q_k in $\mathfrak{C}$ such that $f - g$ is in $\mathfrak{N}$.

Consider an arbitrary $\epsilon > 0$. We may assume ϵ is small enough to ensure that ϵu is interior to $\mathfrak{N}$, using the continuity of scalar multiplication. Choose $\lambda_0, \lambda_1, \ldots, \lambda_N$ with $\lambda_k - \lambda_{k-1} = \epsilon$ for $k = 1, \ldots, N$ and $\lambda_0 u \leqslant f \leqslant \lambda_N u$. For notational simplicity let $f_k = f - \lambda_k u$. By the hypothesis of separation there exists for each k a net $P_k(t)$ in $\mathfrak{C}$ such that

$$P_k f_k^+ \to 0 \qquad \text{and} \qquad P_k' f_k^- \to 0, \tag{4.1}$$

the limits being taken with respect to t. (We hereafter abbreviate $P(t)$ to P.) Since $f_0^- = 0$ we may assume $P_0 = 0$. Also, since $f_N^+ = 0$ we may take the net P_N such that

$$P_N u \to u \tag{4.2}$$

applying the separation hypothesis to u and 0. Now,

$$\begin{aligned} 0 < \epsilon P_{k-1}P'_k u = P_{k-1}P'_k(f_{k-1} - f_k) &< P_{k-1}P'_k(|f_{k-1}| + |f_k|) \\ &< P_{k-1}f^+_{k-1} + P'_k f^-_{k-1} + P_{k-1}f^+_k + P'_k f^-_k < 2P_{k-1}f^+_{k-1} + 2P'_k f^-_k \end{aligned} \tag{4.3}$$

since $f_{k-1}^- < f_k^-$ and $f_k^+ < f_{k-1}^+$. Since the right side of (4.3) converges to 0 by (4.1), we have via (3.1)

$$P_{k-1}P'_k u \to 0. \tag{4.4}$$

Since $P_k P_{k-1}' = (P_k - P_{k-1}) + P_{k-1}P_k'$ and $P_0 = 0$,

$$\sum P_k P'_{k-1} = P_N + \sum P_{k-1}P'_k \tag{4.5}$$

with summation over $k = 1, \ldots, N$. Applying (4.5) to u and taking limits with respect to t, we obtain via (4.4) and (4.2)

$$\sum P_k P'_{k-1} u \to u. \tag{4.6}$$

Recalling that $|f| < \lambda u$ and $P_{k-1}P_k'$ is a contractor, we have

$$|P_{k-1}P_k' f| < \lambda P_{k-1}P_k' u$$

by (2.5) and (2.9). Hence (4.4) gives

$$P_{k-1}P'_k f \to 0. \tag{4.7}$$

Similarly, since $P_N' u \to 0$ by (4.2), $P_N' f \to 0$. So (4.5) and (4.7) give

$$\sum P_k P'_{k-1} f \to f. \tag{4.8}$$

Now since $f_k^- < f_{k-1}^- + \epsilon u$,

$$|P_k P'_{k-1} f_k| < P_k f^+_k + P_k P'_{k-1} f^-_k < P_k f^+_k + P'_{k-1} f^-_{k-1} + \epsilon P_k P'_{k-1} u. \tag{4.9}$$

Thus,

$$\begin{aligned} \left|f - \sum \lambda_k P_k P'_{k-1} u\right| &< \left|f - \sum P_k P'_{k-1} f\right| + \left|\sum P_k P'_{k-1} f_k\right| \\ &< \left|f - \sum P_k P'_{k-1} f\right| + \sum P_k f^+_k \\ &\quad + \sum P'_{k-1} f^-_{k-1} + \epsilon \sum P_k P'_{k-1} u. \end{aligned} \tag{4.10}$$

By (4.8), (4.1), and (4.6) the right side of (4.10) converges to ϵu, which is interior to $\mathfrak{N}$. Hence, the right side of (4.10) is eventually in $\mathfrak{N}$. By (3.1), the left side of (4.10) is likewise eventually in $\mathfrak{N}$, which proves the lemma.

5. Approximation by projectors on a limit unit.

THEOREM 1. *Let $\mathfrak{R}$ be a Boolean ring of projectors on a topological vector lattice $\mathfrak{L}$ and u be a limit unit in $\mathfrak{L}$. Then $\mathfrak{R}u$, the set of all Eu for E in $\mathfrak{R}$, generates $\mathfrak{L}$ if, and only if, $\mathfrak{R}$ separates every pair f and g in $\mathfrak{L}$ for which $f \wedge g = 0$.*

Proof. Let $\mathfrak{R}u$ generate $\mathfrak{L}$. Then, given $f \wedge g = 0$, there exists a net f_ι converging to f and a corresponding net g_ι converging to g of the form:

$$f_\iota = \sum \alpha_k E_k u, \qquad g_\iota = \sum \beta_k E_k u \tag{5.1}$$

where E_k is in $\mathfrak{R}$ and $E_i E_j = 0$ for $i \neq j$. Since $f_\iota \to f$, $|f_\iota| \to |f|$ by (3.1). Moreover $f \geq 0$, so we may assume $f_\iota \geq 0$, and similarly $g_\iota \geq 0$. That is, $\alpha_k \geq 0$ and $\beta_k \geq 0$ in (5.1). Let A_ι be the sum of those E_k in (5.1) for which $\alpha_k \leq \beta_k$. Since $f_\iota \wedge g_\iota = \sum \delta_k E_k u$ where δ_k is the smaller of α_k and β_k, we have $0 \leq A_\iota f_\iota \leq f_\iota \wedge g_\iota$ and $0 \leq A_\iota' g_\iota \leq f_\iota \wedge g_\iota$. Therefore

$$\begin{aligned} A_\iota f &\leq |A_\iota f - A_\iota f_\iota| + A_\iota f_\iota \\ &\leq |f - f_\iota| + f_\iota \wedge g_\iota. \end{aligned} \tag{5.2}$$

Since $f \wedge g = 0$,

$$f_\iota \wedge g_\iota \leq |f \wedge g - f \wedge g_\iota| + |f \wedge g_\iota - f_\iota \wedge g_\iota| \leq |g - g_\iota| + |f - f_\iota|$$

by (2.3). Hence (5.2) gives $|A_\iota f| \leq |g - g_\iota| + 2|f - f_\iota|$. Since $f_\iota \to f$ and $g_\iota \to g$, $A_\iota f \to 0$ by (3.1). Similarly

$$|A_\iota' g| \leq |f - f_\iota| + 2|g - g_\iota|.$$

Hence, $A_\iota' g \to 0$.

The converse follows directly from the fundamental lemma, since PQ' is in $\mathfrak{R}$ for P and Q in $\mathfrak{R}$.

6. Topological lattice algebras. Let $\mathfrak{A}$ be a T_1 topological vector lattice in which an associative, distributive multiplication is defined making $\mathfrak{A}$ a topological algebra with a multiplicative unit 1 which is also a limit unit. Moreover, let $fg \geq 0$ whenever both $f \geq 0$ and $g \geq 0$. We call $\mathfrak{A}$ a *topological lattice algebra*. From **(2)** it follows that multiplication is commutative in $\mathfrak{A}$.

We shall apply the results of the preceding sections by viewing the elements of $\mathfrak{A}$ as operators on $\mathfrak{A}$ via multiplication. This is effective because the operator ordering for elements of $\mathfrak{A}$ is just the ordering in $\mathfrak{A}$. A few simple lemmas serve to establish the basic properties of $\mathfrak{A}$.

LEMMA 1. *If $f \wedge g = 0$, then $fg = 0$.*

Proof. Let $f_n = f \wedge n1$ and $g_n = g \wedge n1$. Since 1 is a limit unit $f_n \to f$ and $g_n \to g$. Since multiplication is continuous $f_n g_n \to fg$. Thus, it suffices to show $f_n g_n = 0$. Since $0 \leq f_n \leq f$ and $0 \leq g_n \leq g$ we have $0 \leq f_n \wedge g_n \leq f \wedge g$. So $f_n \wedge g_n = 0$, since $f \wedge g = 0$. Moreover, $0 \leq f_n \leq n1$ and since $g_n \geq 0$, $0 \leq f_n g_n \leq n g_n$. Similarly $f_n g_n \leq n f_n$. Hence

$$0 \leq \frac{1}{n} f_n g_n \leq f_n \wedge g_n,$$

and so $f_n g_n = 0$.

LEMMA 2. *$f^2 = |f|^2$. Hence, $f^2 \geqslant 0$.*

Proof. By Lemma 1, $f^+ f^- = 0$. So $f^2 = (f^+ - f^-)^2 = f^{+2} + f^{-2} = |f|^2$.

LEMMA 3. *If $f^2 = 0$, then $f = 0$.*

Proof. By Lemma 2 we may assume without loss of generality that $f \geqslant 0$. Consider any $\epsilon > 0$. Now $(f - \epsilon 1)^2 = -2\epsilon f + \epsilon^2 1$, which is positive by Lemma 2. So $2\epsilon f \leqslant \epsilon^2 1$. Dividing by ϵ we get $0 \leqslant 2f \leqslant \epsilon 1$. Letting $\epsilon \to 0$ gives $f = 0$.

LEMMA 4. *If $f \geqslant 0$, $g \geqslant 0$, and $fg = 0$, then $f \wedge g = 0$.*

Proof. Let $h = f \wedge g$. Then $0 \leqslant h \leqslant f$ and $0 \leqslant h \leqslant g$. Therefore $0 \leqslant h^2 \leqslant fh \leqslant fg \leqslant 0$. So $h^2 = 0$. By Lemma 3, $h = 0$.

LEMMA 5. $|fg| = |f|\,|g|$.

Proof. $fg = (f^+ - f^-)(g^+ - g^-) = (f^+g^+ + f^-g^-) - (f^+g^- + f^-g^+)$, a difference of two positive terms. That the product of these two terms is 0 follows from Lemma 1, using the commutative, distributive, and associative laws. Hence, by Lemma 4, the two terms are disjoint. Thus,

$$(fg)^+ = f^+g^+ + f^-g^-$$

and

$$(fg)^- = f^+g^- + f^-g^+.$$

Therefore,

$$|fg| = (fg)^+ + (fg)^- = (f^+ + f^-)(g^+ + g^-) = |f|\,|g|.$$

LEMMA 6. *$fg = 0$ if, and only if, $|f| \wedge |g| = 0$.*

Proof. By Lemma 5, $fg = 0$ if, and only if, $|f|\,|g| = 0$. By Lemmas 1 and 4, $|f|\,|g| = 0$ if, and only if, $|f| \wedge |g| = 0$.

7. Projectors on a topological lattice algebra.

LEMMA 7. *The identity*

$$(Ef)g = f(Eg) = (Ef)(Eg) \tag{7.1}$$

holds for every projector E on $\mathfrak{A}$.

Proof. $(Ef)g - f(Eg) = (Ef)(E'g) - (Eg)(E'f)$, an identity which can be verified by setting $E' = I - E$ on the right and expanding. We shall show that each of the terms on the right side of this identity is 0, in order to derive the first equation in (7.1). Now by (2.9), (2.5), and (2.10),

$$|Ef| \wedge |E'g| = E|f| \wedge E'|g| \leqslant E(|f| + |g|) \wedge E'(|f| + |g|) = EE'(|f| + |g|) = 0.$$

Thus, by Lemma 6, $(Ef)(E'g) = 0$. Similarly $(Eg)(E'f) = 0$. The second equation in (7.1) follows if we replace f in the first equation by Ef.

LEMMA 8. *The projectors E on $\mathfrak{A}$ are isomorphic to the idempotent elements e of $\mathfrak{A}$ via the correspondence $E \sim e$ induced by*

$$E1 = e \tag{7.2}$$

and

$$ef = Ef. \tag{7.3}$$

Proof. Given any idempotent $e = e^2$ in $\mathfrak{A}$, Lemma 2 implies $e \geqslant 0$. Since $1 - e$ is also idempotent we have $0 \leqslant e \leqslant 1$. Thus E defined by (7.3) is a projector. Conversely, every projector E defines an idempotent e via (7.2) which, by Lemma 7, satisfies (7.3). Clearly, $I \sim 1$ and for $A \sim a$ and $B \sim b$, $AB \sim ab$.

The next theorem follows directly from Theorem 1 via Lemmas 6 and 8.

THEOREM 2. *Let $\mathfrak{R}$ be a Boolean ring of idempotents in a topological lattice algebra $\mathfrak{A}$. Then $\mathfrak{R}$ generates $\mathfrak{A}$ if, and only if, $\mathfrak{R}$ separates every pair f and g in $\mathfrak{A}$ for which $fg = 0$.*

8. Subalgebras dense in $\mathfrak{A}$. A *subalgebra* of $\mathfrak{A}$ is a linear subspace which is closed under multiplication.

THEOREM 3. *Let $\mathfrak{R}$ be a subalgebra of a topological lattice algebra $\mathfrak{A}$. Then $\mathfrak{R}$ is dense in $\mathfrak{A}$ if, and only if, $\mathfrak{R}$ separates every pair f and g in $\mathfrak{A}$ for which $fg = 0$.*

To prove this theorem we need another lemma.

LEMMA 9. *The following conditions are equivalent:*

(i) *$\mathfrak{R}$ separates f and g whenever $fg = 0$.*

(ii) *The set of all contractors in the closure of $\mathfrak{R}$ separates f and g whenever $f \wedge g = 0$.*

Proof. We first show that (i) implies that the closure of $\mathfrak{R}$ is a lattice and contains the unit 1. Now the trivial identity $f - g = (f - f \wedge g) - (g - f \wedge g)$ gives, in view of (2.2),

$$(f - g)^+ = f - f \wedge g. \tag{8.1}$$

Thus, to show that the closure of $\mathfrak{R}$ is a lattice we need only show that it contains f^+ whenever it contains f. Since $f^+f^- = 0$, (i) implies the existence of a net h_ι in $\mathfrak{R}$ such that $h_\iota f^+ \to f^+$ and $h_\iota f^- \to 0$. Hence $h_\iota f \to f^+$. Since $h_\iota f$ is in the closure of $\mathfrak{R}$, so is f^+. That 1 is in the closure of $\mathfrak{R}$ follows from (i), since $\mathfrak{R}$ must separate 1 from 0.

Given $f \wedge g = 0$, (i) gives a net h_ι in $\mathfrak{R}$ with $h_\iota f \to 0$ and $h_\iota g \to g$. Let $p_\iota = |h_\iota| \wedge 1$ which is in the closure of $\mathfrak{R}$ by the preceding arguments. Clearly, p_ι is a net of contractors: $0 \leqslant p_\iota \leqslant 1$. Moreover, since $0 \leqslant p_\iota \leqslant |h_\iota|$, $0 \leqslant p_\iota f \leqslant |h_\iota f|$ using Lemma 5. So by (3.1), $p_\iota f \to 0$. From the identity (8.1) we

have $1 - p_\iota = (1 - |h_\iota|)^+$. So $(1 - p_\iota)g \leq |(1 - |h_\iota|)g| \leq |g - h_\iota g|$. Hence, $p_\iota g \to g$. Thus (i) implies (ii).

Given (ii) and $fg = 0$, $|f| \wedge |g| = 0$ by Lemma 6. So there exists a net of contractors p_ι in the closure of $\mathfrak{R}$ separating $|g|$ from $|f|$: $p_\iota|f| \to 0$ and $p_\iota|g| \to |g|$ with $0 \leq p_\iota \leq 1$. Using Lemma 5 we have $p_\iota f \to 0$ and $(1 - p_\iota)g \to 0$. Since p_ι is in the closure of $\mathfrak{R}$ there exists h_ι in $\mathfrak{R}$ such that $p_\iota - h_\iota \to 0$. Hence $|h_\iota f| \leq |h_\iota - p_\iota|\,|f| + p_\iota|f|$ and $|(1 - h_\iota)g| \leq (1 - p_\iota)|g| + |p_\iota - h_\iota|\,|g|$. So $h_\iota f \to 0$ and $h_\iota g \to g$, giving (i).

Proof of Theorem 3. Given (i) we have (ii) by Lemma 9. By the Fundamental Lemma, (ii) implies $\mathfrak{R}$ is dense in $\mathfrak{A}$. Conversely, we shall show that if the closure of $\mathfrak{R}$ is $\mathfrak{A}$, then (ii), and hence (i) holds.

Given $f \wedge g = 0$ let

$$p_n = n\left(g \wedge \frac{1}{n}1\right).$$

We contend that p_n is a sequence of contractors separating g from f. Clearly, $0 \leq p_n \leq 1$. Since $0 \leq p_n \leq ng$, $0 \leq p_nf \leq nfg$. Now $fg = 0$ by Lemma 6, so $p_nf = 0$.

Noting that

$$1 - p_n = n\left(\frac{1}{n}1 - g \wedge \frac{1}{n}1\right),$$

apply (2.2) to $1/n$ 1 and g to obtain, via Lemma 6,

$$(1 - p_n)\left(g - \frac{1}{n}p_n\right) = 0.$$

So

$$(1 - p_n)g = \frac{1}{n}p_n(1 - p_n).$$

Hence,

$$0 \leq (1 - p_n)g \leq \frac{1}{n}1.$$

So $(1 - p_n)g \to 0$.

9. Absolutely continuous set functions. Let u be a bounded, non-negative, finitely additive measure on a Boolean algebra $\mathfrak{B}$ with unit I. The Banach lattice $\mathfrak{V}$ dealt with in **(3)** and **(6)** consists of all finitely additive, real valued functions f on $\mathfrak{B}$ which are absolutely continuous with respect to u:

$$f(E) \to 0 \quad \text{as} \quad u(E) \to 0. \tag{9.1}$$

The norm in $\mathfrak{V}$ is defined by

$$\|f\| = \sup f(E) - f(E') \tag{9.2}$$

where the supremum is taken over all E in $\mathfrak{B}$. The partial ordering is induced by defining $f \geqslant 0$ whenever $f(E) \geqslant 0$ for all E in $\mathfrak{B}$. With this ordering

$$f \wedge g(A) = \inf f(EA) + g(E'A) \tag{9.3}$$

and

$$f \vee g(A) = \sup f(EA) + g(E'A) \tag{9.4}$$

taken over all E in $\mathfrak{B}$ **(1, 3, 4, 6)**. Since $|f| = f \vee -f$, (9.2) and (9.4) give

$$||f|| = |f|\,(I). \tag{9.5}$$

Every E in $\mathfrak{B}$ defines a projector E given by

$$Ef(A) = f(EA) \tag{9.6}$$

for all A in $\mathfrak{B}$. Thus $\mathfrak{B}$, modulo the ideal of all E with $u(E) = 0$, is isomorphic to a subalgebra of the Boolean algebra of all projectors on $\mathfrak{B}$.

Now (9.1) implies that u is a limit unit. To prove this let $f \geqslant 0$ and $f_n = f \wedge nu$. The sequence $(f - f_n)(I)$ is decreasing, hence converges to some limit λ. In view of (9.5) we need only show $\lambda = 0$. By (9.3), $f_n(I) = \inf f(E') + nu(E)$. Hence we may choose a sequence E_n such that

$$f_n(I) \leqslant f(E_n') + n\,u(E_n) \leqslant f_n(I) + \frac{1}{n}.$$

Multiplying by -1 and adding $f(I)$ we obtain

$$(f - f_n)(I) - \frac{1}{n} \leqslant f(E_n) - n\,u(E_n) \leqslant (f - f_n)(I).$$

Hence $f(E_n) - n\,u(E_n)$ converges to λ. Now $0 \leqslant f(E_n) \leqslant f(I)$ and $0 \leqslant \lambda \leqslant f(I)$ while n increases without bound. Hence $u(E_n)$ must converge to 0. By (9.1), $f(E_n)$ does likewise. So $\lambda = -\lim n\,u(E_n)$. Thus $\lambda \leqslant 0$. But $\lambda \geqslant 0$. So $\lambda = 0$.

Given $f \wedge g = 0$ there exists, via (9.3) with $A = I$, a sequence E_n in $\mathfrak{B}$ such that

$$f(E_n) + g(E_n') \to 0. \tag{9.7}$$

By (9.6) and (9.5), $||E_nf|| = f(E_n)$ and $||E_n'g|| = g(E_n')$. So (9.7) implies that $\mathfrak{B}$ separates f and g. By Theorem 1, $\mathfrak{B}u$ generates $\mathfrak{B}$. That is, the "step functions" are dense in $\mathfrak{B}$. (See **(3)** and **(6)**.) As was pointed out by Bochner **(3)**, this gives the Radon-Nikodym theorem **(11)**.

10. The finitely additive integral. Let $\mathfrak{B}$ be a Boolean algebra of subsets E of a set I with I as unit. Let u be a bounded, non-negative, finitely additive measure on $\mathfrak{B}$. A *partition* Δ is a finite class of disjoint sets in $\mathfrak{B}$ whose union is I. The partitions are ordered by defining $\Delta' \geqslant \Delta$ whenever Δ' is a refinement of Δ. For $f(x)$ real-valued on the domain I and $\Delta = \{E_1, \ldots, E_n\}$ any partition, let

$$s(\Delta) = \sum f(x_k)u(E_k) \tag{10.1}$$

where x_k is any point in E_k and k ranges through $1, \ldots, n$. In general, $s(\Delta)$ is a many-valued function of Δ, a particular value depending on the choice of x_k in E_k. If $\lim s(\Delta)$ exists (in the Moore-Smith sense **(8)**) uniformly for all such choices, then f is said to be *integrable*.

Introducing the upper and lower Darboux sums

$$\bar{s}(\Delta) = \sum \sup f(x_k) u(E_k) \tag{10.2}$$

and

$$\underline{s}(\Delta) = \sum \inf f(x_k) u(E_k),$$

let $S(\Delta, f) = \bar{s}(\Delta) - \underline{s}(\Delta)$. In (10.2) we assume $\infty \,.\, 0 = 0$. Since $\lim\sup s(\Delta) = \lim \bar{s}(\Delta)$ and $\lim\inf s(\Delta) = \lim \underline{s}(\Delta)$, f is integrable if, and only if, $\lim S(\Delta, f) = 0$. Note that for any f, $S(\Delta, f)$ is a decreasing function of Δ. Since $S(\Delta, \alpha f + \beta g) \leqslant |\alpha|\, S(\Delta, f) + |\beta|\, S(\Delta, g)$ the integrable functions form a vector space. Since $S(\Delta, 1) = 0$ the constant functions are integrable. That products of integrable functions are integrable follows from the inequality $S(\Delta, fg) \leqslant M(f)\, S(\Delta, g) + M(g)\, S(\Delta, f)$ where $M(f)$ is the supremum of $|f(x)|$ for x restricted to those sets in Δ which are not of measure zero. That $|f|$ is integrable whenever f is integrable follows from the inequality $S(\Delta, |f|) \leqslant S(\Delta, f)$. Given $|f(x) - g(x)| < \epsilon$ for all x we have $S(\Delta, f) \leqslant S(\Delta, g) + S(\Delta, f - g) \leqslant S(\Delta, g) + 2\epsilon u(I)$. So a uniform limit of integrable functions is integrable. Since an integrable function is bounded except on a set of measure zero, we shall consider only bounded integrable functions. These form a topological lattice algebra under uniform convergence with the usual ordering and algebraic operations. Using Theorem 2, we shall show that this algebra is generated by its idempotents. Thus, it suffices to show that for f any bounded integrable function, f^- can be separated from f^+ by integrable idempotents.

Consider any $\epsilon > 0$. Choose a sequence Δ_n of partitions such that $\Delta_{n+1} \geqslant \Delta_n$ and $S(\Delta_n, f) \to 0$, which is possible because f is integrable. Let C_n be the union of those sets E, belonging to the partition Δ_n, for which there exist x and y in E with $f^+(x) \geqslant \epsilon$ and $f^-(y) \geqslant \epsilon$. By induction, starting with $A_0 = B_0 = \phi$ and $C_0 = I$, let A_n be the union of A_{n-1} and those sets E in Δ_n which are contained in C_{n-1} and have $f^+(x) < \epsilon$ for all x in E. Let B_n be the union of B_{n-1} and those sets E in Δ_n which are contained in C_{n-1}, have $f^-(x) < \epsilon$ for all x in E, and have $f^+(y) \geqslant \epsilon$ for some y in E. Then A_{n-1} is a subset of A_n, B_{n-1} of B_n, and C_n of C_{n-1}. Since $2\epsilon u(C_n) \leqslant S(\Delta_n, f)$, we have $u(C_n) \to 0$. Let $A = \lim A_n$ and $C = \lim C_n$. Let E be the union of A with the set of all points x in C for which $f^+(x) = 0$. Let e be the indicator of E:

$$e(x) = \begin{cases} 1 \text{ for } x \text{ in } E \\ 0 \text{ for } x \text{ in } E'. \end{cases} \tag{10.3}$$

Since A_n is contained in E and B_n is contained in E', $e(x)$ equals 1 for x in A_n and 0 for x in B_n. Hence, $S(\Delta_n, e) \leqslant u(C_n)$ which converges to 0. So e

is integrable. For x in E either x is in C with $f^+(x) = 0$ or x belongs to some A_n, implying $f^+(x) < \epsilon$. Clearly then $ef^+ < \epsilon 1$. For x in E', either x is in C with $f^+(x) > 0$, hence $f^-(x) = 0$, or x is in some B_n, implying $f^-(x) < \epsilon$. So $(1 - e)f^- < \epsilon 1$.

Thus, by Theorem 2, the algebra of bounded integrable functions is generated under uniform convergence by its idempotents.

A similar result can be obtained for the almost everywhere continuous functions on a closed interval, using Theorem 2. Combining these two results, we get Lebesgue's characterization of the Riemann integrable functions **(7)**.

References

1. G. Birkhoff, *Lattice theory*, A.M.S. Coll. Pub. (New York, 1940).
2. G. Birkhoff and R. S. Pierce, *Lattice-ordered rings*, Anais da Acad. Brasileira de Ciencias, *28* (1956), 41–69.
3. S. Bochner, *Additive set functions on groups*, Ann. Math., *40* (1939), 769–99.
4. S. Bochner and R. S. Phillips, *Additive set functions and vector lattices*, Ann. Math., *42* (1941), 316–24.
5. H. Freudenthal, *Teilweise geordnete Moduln*, Proc. Acad. Wet. Amsterdam, *39* (1936), 641–51.
6. S. Leader, *The theory of L^p-spaces for finitely additive set functions*, Ann. Math., *58* (1953), 528–43.
7. H. Lebesgue, *Lecons sur l'intégration et la recherche des fonctions primitives*, Gauthier-Villars (Paris, 1928).
8. E. H. Moore and H. L. Smith, *A general theory of limits*, Amer. J. Math., *44* (1922), 102–21.
9. H. Nakano, *Modern spectral theory* (Tokyo, 1950).
10. I. Namioka, *Partially ordered linear topological spaces*, Amer. Math. Soc., Mem. 24 (1957).
11. S. Saks, *Theory of the integral* (Warsaw, 1937).
12. M. H. Stone, *Applications of the theory of boolean rings to general topology*, Trans. Amer. Math. Soc. *41* (1937), 375–481.

Rutgers University

TENSOR PRODUCTS OF BANACH ALGEBRAS

BERNARD R. GELBAUM[1]

1. Introduction. This paper is concerned with a generalization of some recent theorems of Hausner **(1)** and Johnson **(4; 5)**. Their result can be summarized as follows: *Let G be a locally compact abelian group, A a commutative Banach algebra, $B^1 = B^1(G, A)$ the (commutative Banach) algebra of A-valued, Bochner integrable functions on G, $\mathfrak{M}_1$ the maximal ideal space of A, $\mathfrak{M}_2$ the maximal ideal space of $L^1(G)$ (the [commutative Banach] algebra of complex-valued, Haar integrable functions on G), $\mathfrak{M}_3$ the maximal ideal space of B^1. Then $\mathfrak{M}_3$ and the Cartesian product $\mathfrak{M}_1 \times \mathfrak{M}_2$ are homeomorphic when the spaces $\mathfrak{M}_i$, $i = 1, 2, 3$, are given their weak* topologies. Furthermore, the association between $\mathfrak{M}_3$ and $\mathfrak{M}_1 \times \mathfrak{M}_2$ is such as to permit a description of any epimorphism $E_3: B^1 \to B^1/M_3$ in terms of related epimorphisms $E_1: A \to A/M_1$ and $E_2: L^1(G) \to L^1(G)/M_2$, where M_i is in $\mathfrak{M}_i$, $i = 1, 2, 3$.*

On the other hand, Hausner **(2)** (and the author, independently) showed that a similar result is valid for generalized continuous function algebras. One form of the theorem is the following: *Let X be a compact Hausdorff space, A a commutative Banach algebra, $D = C(X, A)$ the (commutative Banach) algebra of A-valued continuous functions on X, $\mathfrak{M}_1$ the maximal ideal space of A, $\mathfrak{M}_2$ the maximal ideal space of $C(X)$ (the [commutative Banach] algebra of complex-valued continuous functions on X), $\mathfrak{M}_3$ the maximal ideal space of D. Then $\mathfrak{M}_3$ and the Cartesian product $\mathfrak{M}_1 \times \mathfrak{M}_2$ are homeomorphic when the spaces $\mathfrak{M}_i$, $i = 1, 2, 3$, are given their weak* topologies. Furthermore, the association between $\mathfrak{M}_3$ and $\mathfrak{M}_1 \times \mathfrak{M}_2$ is such as to permit a description of any epimorphism $E_3: D \to D/M_3$ in terms of related epimorphisms $E_1: A \to A/M_1$ and $E_2: C(X) \to C(X)/M_2$, where M_i is in $\mathfrak{M}_i$, $i = 1, 2, 3$.*

The crucial point in the latter theorem is the proof that D is spanned by "simple" functions, that is, functions which are linear combinations, with coefficients in A, of complex-valued continuous functions on X. On the other hand, the very definition of B^1 shows that it is spanned by "simple" functions, that is, this time, functions which are linear combinations, with coefficients in A, of complex-valued, Haar integrable functions on G. Clearly, in each instance, the collection of "simple" functions is an algebra which is a tensor

Received May 6, 1958. This research was supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command, under contract No. AF 49 (638)-64. Reproduction in whole or in part is permitted for any purpose of the United States Government.

[1]The author is indebted to Professor G. K. Kalisch for many stimulating conversations on the subject matter of this investigation.

product of A and some complex function algebra, and the object of discussion is the completion of this tensor product with respect to an appropriate norm.

2. Tensor products. Let A_1 and A_2 be Banach algebras and let $A_3' = A_1 \otimes A_2$ be their algebraic tensor product **(0)**. As is well known, **(6)** there are generally many norms which can be given to A_3' in terms of the norms of A_1 and A_2. Our first result is about one of these norms.

THEOREM 1. *Let* $\|\ldots\|_i$ *be the norms in* A_i, $i = 1, 2$. *Then the "greatest cross norm"* **(6)** *defined in* A_3' *by*

$$\left\|\sum_{i=1}^{n} a_1^{(i)} \otimes a_2^{(i)}\right\|_3' = \inf \sum_{i=1}^{n} \|a_1^{(i)}\|_1 \|a_2^{(i)}\|_2,$$

where the inf *is taken over the equivalence class which defines*

$$\sum_{i=1}^{n} a_1^{(i)} \otimes a_2^{(i)},$$

is a Banach algebra norm which satisfies the relationship

$$\mathfrak{P}' \qquad \| a_1 \otimes a_2 \|_3' = \| a_1 \|_1 \| a_2 \|_2.$$

Proof. In **(6)** the validity of the last equality is shown. We prove here the fact that if p, q are in A_3', then $\|pq\|_3' < \|p\|_3'\|q\|_3'$. To this end, let $r > 0$ be given. Then there is a choice of $a_1^{(i)}$, $a_2^{(i)}$, $b_1^{(j)}$, $b_2^{(j)}$ for which

$$\sum_{i=1}^{n} a_1^{(i)} \otimes a_2^{(i)} \quad \text{and} \quad \sum_{j=1}^{m} b_1^{(j)} \otimes b_2^{(j)}$$

define the respective equivalence classes of p and q and for which

$$\|p\|_3'\|q\|_3' > \left(\sum_{i=1}^{n} \|a_1^{(i)}\|_1\|a_2^{(i)}\|_2\right)\left(\sum_{j=1}^{m} \|b_1^{(j)}\|_1\|b_2^{(j)}\|_2\right) - r.$$

The last expression is

$$\sum_{i,j} \|a_1^{(i)}\|_1\|b_1^{(j)}\|_1\|a_2^{(i)}\|_2\|b_2^{(j)}\|_2 - r$$

which majorizes

$$\sum_{i,j} \|a_1^{(i)}b_1^{(j)}\|_1\|a_2^{(i)}b_2^{(j)}\|_2 - r.$$

Obviously, the last expression majorizes $\|pq\|_3' - r$. Since r is an arbitrary positive number, we see $\|p\|_3'\|q\|_3' > \|pq\|_3'$. This completes the proof.

THEOREM 2. *Let* A_3 *be the completion of* A_3' *endowed with the "greatest cross norm"* $\|\ldots\|_3'$. *Let* A_i *be commutative and let* $\mathfrak{M}_i$ *be their respective maximal ideal spaces with their respective weak* topologies,* $i = 1, 2$. *Then* A_3 *is a commutative Banach algebra. Its norm* $\|\ldots\|_3$ *satisfies the analogue*

$$\mathfrak{P} \qquad \|a_1 \otimes a_2\|_3 = \|a_1\|_1\|a_2\|_2$$

of $\mathfrak{P}'$. If $\|\ldots\|_3''$ is any tensor product norm relative to which A_3' is a normed algebra, with no dense reg. max. ideal (for example, if $\|\cdots\|_3''$ is the greatest cross norm), and if A_3 is the completion of A_3' relative to $\|\cdots\|_3''$, then A_3 is a commutative Banach algebra and its maximal ideal space $\mathfrak{M}_3$ in its weak topology is homeomorphic with the Cartesian product $\mathfrak{M}_1 \times \mathfrak{M}_2$. Let t be the homeomorphism the existence of which is asserted: $t: \mathfrak{M}_3 \to \mathfrak{M}_1 \times \mathfrak{M}_2$, and let $t(M_3) = (M_1, M_2)$. Then the epimorphisms*

$$E_1: A_1 \to A_1/M_1, \quad E_2: A_2 \to A_2/M_2, \quad E_3: A_3 \to A_3/M_3,$$

are uniquely determined by the respective maximal ideals M_i, $i = 1, 2, 3$; E_1 and E_2 together determine E_3 and conversely.

Proof. The commutativity of A_3 and the validity of $\mathfrak{P}$ are clear consequences of the hypotheses.

Since A_i/M_i, $i = 1, 2, 3$, is the complex numbers, and since each epimorphism E_i, $i = 1, 2, 3$ commutes with multiplication by complex numbers ($cE_i(a) = E_i(ca)$, c complex, a in A_i) and since the complex numbers admit no non-trivial automorphism which commutes with multiplication by complex numbers, the uniqueness of the E_i follows.

We now proceed to set up a $1-1$ correspondence between $\mathfrak{M}_3$ and $\mathfrak{M}_1 \times \mathfrak{M}_2$. After this has been accomplished, the correspondence will be shown to be a homeomorphism. With a view to greater ultimate generality, we shall, however, show how to establish the kind of correspondence we need between $\mathfrak{M}_1 \times \mathfrak{M}_2$ and a part of $\mathfrak{M}_3$ under conditions far less restrictive than those imposed in the hypothesis of Theorem 2. This correspondence will serve when the hypothesis of Theorem 2 is in force and will in fact prove to be the homeomorphism which is sought. What follows then is an interlude, justified and required by economy.

During this interlude we shall not assume that A_1 and A_2 are commutative. $\mathfrak{M}_1$ and $\mathfrak{M}_2$ will denote their respective spaces of (two-sided) regular maximal ideals. For each pair (M_1, M_2) in $\mathfrak{M}_1 \times \mathfrak{M}_2$, let E_1 and E_2 be *some* epimorphisms $E_i: A_i \to A_i/M_i$, $i = 1, 2$. For p in A_3', define E_3' by the formula

$$E_3'(p) = \sum_{i=1}^{n} E_1(a_1^{(i)}) \otimes E_2(a_2^{(i)}),$$

a member of the tensor product $(A_1/M_1) \otimes (A_2/M_2)$, where

$$\sum_{i=1}^{n} a_1^{(i)} \otimes a_2^{(i)}$$

is some representation of p. Clearly $E_3'(p)$ does not depend on the representation of p and is an epimorphism of A_3'. $E_3': A_3' \to (A_1/M_1) \otimes (A_2/M_2)$. Let A_3' have the norm $\|\cdots\|''$ and let $E_3'(A_3')$ have the quotient space norm (which is independent of the choice of E_1 and E_2). The quotient space norm is admissible as a true norm since A_3' has no dense reg. max. ideal. Then,

relative to these topologies, E_3' is a bounded (hence uniformly continuous) transformation of A_3' and has a unique extension E_3, an epimorphism of A_3 onto the completion of $(A_1/M_1) \otimes (A_2/M_2)$ relative to its (quotient space) norm.

Let $M_3 = E_3^{-1}(0)$. M_3 is an ideal in A_3. If u_i are identities modulo M_i in A_i, $(i = 1, 2)$, then $E_3(u_1 \otimes u_2)$ is an identity in $E_3(A_3)$, whence M_3 is regular. Consequently, there is a regular maximal ideal N_3 which contains M_3. We shall show that N_3 and M_3 are the same.

For this purpose, we define two mappings G_i of A_i into $E_3(A_3)$, $i = 1, 2$, as follows:

$$G_i(a_i) = E_3(a_i u), \qquad i = 1, 2,$$

where a_i is in A_i and u is an identity modulo M_3. Clearly $G_i(a_i)$ is independent of the choice of u. Let F_3 be *some* epimorphism, $F_3: A_3 \to A_3/N_3$, and let H_i, $i = 1, 2$, be engendered by F_3 as G_i are engendered by E_3. We will show that $M_i = H_i^{-1}(0) = G_i^{-1}(0) = E_i^{-1}(0)$, $(i = 1, 2)$.

If a_1 is in M_1, then $E_1(a_1) = 0$, whence $E_3(a_1 u) \equiv E_3(a_1 u_1 \otimes u_2) = E_1(a_1 u_1) \otimes E_2(u_2) = 0$ (where $u = u_1 \otimes u_2$). Thus $G_1(a_1) = 0$ and hence M_1 is contained in $G_1^{-1}(0)$. Since $G_1^{-1}(0)$ is a proper ideal and M_1 is a maximal ideal we see that $M_1 = G_1^{-1}(0)$.

Since $a_1 u_1 \otimes u_2$ is a member of M_3 which is a subset of N_3, it follows that $F_3(a_1 u_1 \otimes u_2) = 0 = H_1(a_1)$. We see that M_1 is contained in $H_1^{-1}(0)$ which is a proper ideal of A_1. Since M_1 is maximal, it follows that $M_1 = H_1^{-1}(0)$. Of course, by definition, $M_1 = E_1^{-1}(0)$. Analogously, we can show $M_2 = H_2^{-1}(0) = G_2^{-1}(0) = E_2^{-1}(0)$.

In order to continue we shall require the following lemmas.

LEMMA 1. *Let A be a Banach algebra, I a closed ideal of A and let E and E'' be two epimorphisms, $E: A \to A/I$, $E'': A \to A/I$. Then, relative to the quotient space (norm) topology of A/I, there is an isometric automorphism α of A/I, α commutes with complex multiplication and $E = \alpha E''$.*

Proof. For b in A/I, let $E''(a) = b$ and let $\alpha(b) = E(a)$. If $E''(a') = b$, then $a' - a$ is in I, whence $E(a') = E(a)$ and thus $\alpha(b)$ is uniquely defined. If $E(a'') = b$, then $\alpha E''(a'') = E(a'') = b$, whence α is an automorphism, which clearly commutes with complex multiplication. Finally, if $\|\cdots\|_A$ and $\|\cdots\|$ are the respective norms of A and A/I, we see

$$\|\alpha(b)\| = \|E(a)\| = \inf\{\|a + i\|_A \mid i \text{ in } I\}.$$

On the other hand,

$$\|b\| = \|E''(a)\| = \inf\{\|a + i\|_A \mid i \text{ in } I\},$$

whence $\|\alpha(b)\| = \|b\|$ and α is an isometry.

LEMMA 2. *Let A_i be Banach algebras, I_i closed ideals in A_i, E_i, E_i'' epimorphisms, $E_i: A_i \to A_i/I_i$, $E_i'': A_i \to A_i/I_i$, α_i the isometric automorphisms*

(*Lemma* 1) *for which* $E_i = \alpha_i E_i''$, $i = 1, 2$, *and let* α *be the tensor product* $\alpha_1 \otimes \alpha_2$. *If* $A_1 \otimes A_2$ *is given some tensor product norm with respect to which* $A_1 \otimes A_2$ *becomes a normed algebra, then* α *is an isometric automorphism of* $(A_1/I_1) \otimes (A_2/I_2)$ *relative to the quotient space norm described earlier. If* E' *and* $(E')''$ *are the respective epimorphisms engendered by* E_1, E_2 *and* E_1'', E_2'', *then* $E' = \alpha(E')''$.

Proof. The fact that α is an automorphism is clear, as is the relationship $E' = \alpha(E')''$. If b is in $(A_1/I_1) \otimes (A_2/I_2)$, then a representative of b is an expression of the form

$$\sum_{i=1}^{n} E_1''(a_1^{(i)}) \otimes E_2''(a_2^{(i)}).$$

A representative of $\alpha(b)$ is the expression

$$\sum_{i=1}^{n} E_1(a_1^{(i)}) \otimes E_2(a_2^{(i)}).$$

The argument given in Lemma 1 may be repeated *mutatis mutandis* to show that $\alpha(b)$ and b have the same norm.

LEMMA 3. *Let* A *be a normed algebra and let* α *be an isometric automorphism of* A *which commutes with complex multiplication. If the completion* $\bar{A}$ *of* A *is simple, so is the completion* $\overline{\alpha A}$ *of* αA.

Proof. Since α is an isometry, it may be extended in a unique fashion to an isometric automorphism $\bar{\alpha}$ of $\bar{A}$ which commutes with complex multiplication. Clearly $\bar{\alpha}(\bar{A}) = \overline{\alpha A}$, whence the simplicity of $\bar{A}$ implies the simplicity of $\overline{\alpha A}$.

From the preceding paragraphs and lemmas we can conclude that there exist isometric automorphisms α_i, β_i of A_i/M_i which commute with complex multiplication and which satisfy the relations $H_i = \alpha_i G_i = \beta_i E_i$, $(i = 1, 2)$. If β is the tensor product $\beta_1 \otimes \beta_2$, then β is an isometry of (A_1/M_1) and the following relationship is valid:

$$F_3(A_3') = \beta((A_1/M_1) \otimes (A_2/M_2)) = \beta E_3(A_3').$$

Since the completion of $F_3(A_3')$ is $F_3(A_3)$ which is simple, and since β^{-1} is an isometry, we see (Lemma 3) that the completion of $\beta^{-1}F_3(A_3')$ is simple and hence that the completion of $E_3(A_3')$ is simple. Hence $M_3 = E_3^{-1}(0)$ is a regular maximal ideal, and thus $M_3 = N_3$.

We have shown how to associate with each pair (M_1, M_2) *a* maximal ideal M_3. The method of association demands that we show that M_3 is uniquely determined in this manner by the pair (M_1, M_2), regardless of which epimorphisms E_1, E_2, etc., are used in the construction.

To this end, suppose that E_1'', E_2'' are chosen in place of E_1, E_2, at the beginning of our construction. Then there are automorphisms γ_i of A_i/M_i such that $E_i'' = \gamma_i E_i$, $(i = 1, 2)$. Let E_3'' be engendered by E_1'', E_2'' as E_3

is engendered by E_1, E_2, and let E_3'' engender G_1'', G_2'' as E_3 engenders G_1, G_2. Then there are automorphisms π_i of A_i/M_i which satisfy $G_i'' = \pi_i E_i$, $i = 1, 2$. If we set π equal to the tensor product $\pi_1 \otimes \pi_2$, we see that $E_3'' = \pi E_3$ and hence that $(E_3'')^{-1}(0) = E_3^{-1}(0)$. Thus M_3 is uniquely determined, even though the epimorphisms involved in its determination are not unique.

The interlude is over and we continue the proof by using the complete hypothesis of our theorem.

We proceed to establish a correspondence between elements of $\mathfrak{M}_3$ and elements of $\mathfrak{M}_1 \times \mathfrak{M}_2$. If M_3 is in $\mathfrak{M}_3$ and if E_3 is the (unique) epimorphism, $E_3: A_3 \to A_3/M_3$, we can define G_i, $i = 1, 2$, as we did above in the more general context. This time we define M_i to be $G_i^{-1}(0)$, $(i = 1, 2)$. We will show that M_1 and M_2 are maximal ideals which engender, in the manner described above, a maximal ideal which is precisely M_3. The circle will thereby be closed.

The commutativity of A_1 and A_2 implies that A_3' (and hence A_3) is commutative. Let u be an identity modulo M_3. Then if p (in A_3'), represented by

$$\sum_{i=1}^{n} a_1^{(i)} \otimes a_2^{(i)},$$

is so near to u that $E_3(p) \neq 0$, we see that some term in the representation of $E_3(p)$ is not zero. Hence for some i_0,

$$G_1(a_1^{(i_0)}),\ G_2(a_2^{(i_0)})$$

are both not zero. It follows, since A_3/M_3 is the complex number system, that G_i are non-trivial epimorphisms, $G_i: A_i \to C$ (the complex number system), whence M_i are maximal ideals, $(i = 1, 2)$.

Clearly, if M_3'' is the maximal ideal engendered by the M_i in the manner described earlier, then M_3'' contains M_3, and, since M_3 is maximal, M_3'' and M_3 are the same.

Thus we have established a $1-1$ correspondence t between $\mathfrak{M}_3$ and $\mathfrak{M}_1 \times \mathfrak{M}_2$.

The homeomorphism between $\mathfrak{M}_3$ and $\mathfrak{M}_1 \times \mathfrak{M}_2$ can be established as follows. If a is in a commutative Banach algebra A, M is a maximal ideal of A, then $a^+(M)$ denotes the complex number into which a is mapped when A is reduced modulo M. If M_{03} is in $\mathfrak{M}_3$, if $t(M_{03}) = (M_{01}, M_{02})$ and if $N(M_{01}, M_{02})$ is a neighbourhood of (M_{01}, M_{02}) in $\mathfrak{M}_1 \times \mathfrak{M}_2$ we may assume $N(M_{01}, M_{02})$ is of the form $N(M_{01}) \times N(M_{02})$ where $N(M_{0i})$ are neighbourhoods of M_{0i} in M_i, $(i = 1, 2)$. But

$$N(M_{0i}) = \{M_i\, |a_{ji}^+(M_i) - a_{ji}^+(M_{0i})| < r_i, j = 1, 2, \dots, J_i, r_i > 0\}.$$

Consider

$$\begin{aligned} &N(M_{03}) = \\ &\{M_3 |(a_{j1} \otimes u_2)^+(M_3) - (a_{j1} \otimes u_2)^+(M_{03})| < r_1, j = 1, 2, \dots, J_1, \\ &\qquad |(u_1 \otimes a_{j2})^+(M_3) - (u_1 \otimes a_{j2})^+(M_{03})| < r_2, j = 1, 2, \dots, J_2\}, \end{aligned}$$

where u_i are identities modulo M_i, $i = 1, 2$, and $t(M_3) = (M_1, M_2)$. Since $(u_1 \otimes a_2)^+(M_3) = a_2^+(M_2)$, we see $t(N(M_{03}))$ is contained in $N(M_{01}, M_{02})$.

On the other hand, let

$$N(M_{03}) = \{M_3|a_j(M_3) - a_j(M_{03})| < r, j = 1, 2, \dots J\}.$$

Choose

$$P_j = \sum_{i=1}^{n_j} a_{i1}^{(j)} \otimes a_{i2}^{(j)} \quad \text{in} \quad A_3'$$

so that $||a_j - P_j||_3'' < r/3$, $j = 1, 2, \dots, J$. Let

$$N(M_{01}) = \{M_1\,|a_{i1}^{(j)+}(M_1) - a_{i1}^{(j)+}(M_{01})| < r/(6Jn(2R_1 + r)),$$
$$i = 1, 2, \dots, n_j, j = 1, 2, \dots, J\},$$

where

$$n = \sum_{j=1}^{J} n_j,\ R_1 = \sup_{i,j}\{||a_{i1}^{(j)}||_1\}.$$

Similarly let

$$N(M_{02}) = \{M_2|\,|a_{i2}^{(j)+}(M_2) - a_{i2}^{(j)+}(M_{02})| < r/6Jn(2R_2 + r)),$$
$$i = 1, 2, \dots, n_j, j = 1, 2, \dots, J\}.$$

Then if $M_3 = (M_1, M_2)$ is in $N(M_{01}) \times N(M_{02})$, we see

$$|a_j^+(M_3) - a_j^+(M_0)|$$
$$< |(a_j - P_j)^+(M_3)| + |(a_j - P_j)^+(M_{03})| + |P_j^+(M) - P_j^+(M_{03})|$$
$$< 2r/3 + \left| \sum_{i=1}^{n_j} a_{i1}^{(j)+}(M_1)a_{i2}^{(j)+}(M_2) - a_{i1}^{(j)+}(M_{01})a_{i2}^{(j)+}(M_{02}) \right| < r,$$

and hence $t^{-1}(N(M_{01}) \times N(M_{02})) \subset N(M_{03})$, and t is a homeomorphism. The proof of Theorem 2 is complete.

The following remarks are in order at this point.*

1. A little reflection shows that $B^1(G, A)$ is the completion of the tensor product of $L^1(G)$ and A relative to the norm:

$$\left|\left| \sum_{i=1}^{n} \lambda_i(x)a_i \right|\right|' = \int_G \left|\left| \sum_{i=1}^{n} \lambda_i(x)a_i \right|\right|_A dx.$$

2. The result of Hausner and the author shows that $C(X, A)$ is the completion of the tensor product of $C(X)$ and A relative to the norm:

$$\left|\left| \sum_{i=1}^{n} \lambda_i(x)a_i \right|\right|' = \sup\left\{\left|\left| \sum_{i=1}^{n} \lambda_i(x)a_i \right|\right|_A |x \in X\right\}.$$

*At the time of the writing of this paper, the author was unaware of the results of Willcox **(8)** and of the appearance of Hausner's paper **(3)**. Clearly the spirit expressed in the second paragraph, p. 876 of **(8)** has motivated much of our study.

3. The tensorial approach explains and unifies a collection of phenomena and symmetries hitherto observed without comprehension.

For example, Johnson **(5)** shows that $B^1(G, L^1(H))$ and $L^1(G \times H)$ are isomorphic if G and H are locally compact abelian groups. From our standpoint $B^1(G, L^1(H))$ is the completion of the tensor product $L^1(G) \otimes L^1(H)$. Relative to this format, Johnson's theorem is essentially the statement that $L^1(G) \otimes L^1(H)$ (completed) and $L^1(G \times H)$ are isomorphic. The symmetry and truth of this statement are clarified by the tensorial viewpoint.

4. When either of A_1 or A_2 is non-commutative, the most important topologies for the associated spaces of two-sided regular maximal ideals are the kernel-hull topologies **(7)**. In general, under these circumstances, A will be non-commutative and even if the "natural" 1–1 map t: $\mathfrak{M}_3 \to \mathfrak{M}_1 \times \mathfrak{M}_2$ can be constructed, the question of the bi-continuity of t seems to be open.

5. The property $\mathfrak{P}$ of $\|\cdots\|_3$ is irrelevant to the existence of the homeomorphism t. The impact of Theorem 1 and the associated part of Theorem 2 is the existence of norms for A_3' and A_3 relative to which they become normed or Banach algebras.

6. When A_1 and A_2 are not assumed to be commutative, the following results obtain:

(i) *If A_1 and A_2 have "approximate identities," then the* 1–1 *mapping t: $\mathfrak{M}_3 \to \mathfrak{M}_1 \times \mathfrak{M}_2$ can be constructed.*

(ii) *If A_1 and A_2 have identities e_1 and e_2, and if $t(M_3) = (M_1, M_2)$, then $M_2 = M_3 \cap (e_1 \otimes A_2)$ and $M_1 = M_3 \cap (A_1 \otimes e_2)$.*

Proof. Ad(i) By an "approximate identity" in A_i is meant an A_i-valued function v_{ip}, on a directed set P_i such that

$$\lim_{P_i} v_{ip} a_i = a_i$$

for any a_i in A_i, $(i = 1, 2)$. We have observed that $\mathfrak{M}_1 \times \mathfrak{M}_2$ is always naturally embedded in $\mathfrak{M}_3$. On the other hand, for a given M_3 in $\mathfrak{M}_3$ the construction of the naturally associated pair (M_1, M_2) can begin with the mappings G_1, G_2 as above. This time, however, the proof of the regularity and maximality of the relevant ideals proceeds differently.

First, recognizing that A_3 is an A_i-module, we remark that

$$\lim_{P_i} v_{ip}\, q = q, \qquad i = 1, 2,$$

for any q in A_3. Thus, if u is an identity modulo M_3,

$$\lim_{P_i} E_3(v_{ip} u) = E_3(u) = e,$$

the identity of A_3/M_3. This means that $G_i(A_i)$ has e as a point of closure. On the other hand, $G_i(A_i)$ is a complete normed space, and thus e is in $G_i(A_i)$, whence $G_i^{-1}(0) \equiv M_i$ is regular, $i = 1, 2$. The maximality of M_i can be established as in the previous case, once the regularity is known. Of course, our observations on the ambiguity of the epimorphisms can be repeated.

Ad (ii) The existence of t is assured by i. $e_1 \otimes A_2$ and A_2 are isomorphic. Clearly $M_3 \cap (e_1 \otimes A_2)$ is isomorphic to some ideal N_2 in A_2. Let E_3, G_2 have meanings as given earlier. Then, since $e_1 \otimes e_2$ is an identity modulo M_3,

$$G_2(a_2) = E_3((e_1 \otimes e_2)a_2) = E_3(e_1 \otimes a_2)$$

and $G_2(a_2) = 0$ if and only if $E_3(e_1 \otimes a_2) = 0$, that is, if and only if $e_1 \otimes a_2$ is in M_3. Since $e_1 \otimes a_2$ is in $e_1 \otimes A_2$ we see $G_2(a_2) = 0$ if and only if $e_1 \otimes a_2$ is in $M_3 \cap (e_1 \otimes A_2) = N_2$. Thus $N_2 = M_2 = G_2^{-1}(0)$.

3. Group Representations. In the particular case where $A_1 = L^1(G)$, G is a locally compact abelian group, and A_2 is a commutative Banach algebra with an involution and an identity, there are some interesting group representations which can be found.

If $\alpha(x)$ is in G^+ (the character group of G), then for $f(x)$ in $A_3 = B^1(G, A_2)$, the mapping $\pi_\alpha: A_3 \to A_2$ defined by

$$\pi_\alpha(f(x)) = \int_G f(x)\overline{\alpha(x)}\, dx$$

is a homomorphism. If we define an "involution" in A_3 by the formula $f^\dagger(x) = (f(x^{-1}))^*$ where * is the involution in A_2, then $\pi_\alpha(f^\dagger(x)) = (\pi_\alpha(f(x)))^*$. Clearly π_α is continuous, and actually π_α is an epimorphism (which commutes with multiplication by elements of A_2), since $\pi_\alpha(\lambda(x)e_1) = e_1$, if $\lambda(x)$ is in $L^1(G)$ and $\lambda^+(\alpha) = 1$.

On the other hand, let $\mathfrak{J}$ be the non-empty set of inverses in A_2 and let π be a $\dagger_*A_2$-epimorphism: $\pi: A_3 \to A_2$, that is, π commutes with multiplication by elements of A_2 and $\pi(f^\dagger) = (\pi f)^*$. For arbitrary f in $\pi^{-1}(\mathfrak{J})$ define $\alpha_\pi(x)$ by the formula $(\pi(f_x))(\pi f)^{-1}$. Then, in the usual fashion, one can show: $\alpha_\pi(xy) = \alpha_\pi(x)\alpha_\pi(y)$; $\alpha_\pi(x^{-1}) = (\alpha_\pi(x))^*$; $\alpha_\pi(e) = e_2$; $\alpha_\pi(x)$ is f-free, bounded, continuous; $\pi(u_x) \to \alpha_\pi(x)$ for any approximate identity $\{u\}$ in $L^1(G)$, where e is the identity of G, e_2 is the identity of A. We call $\alpha_\pi(x)$ a *unitary representation* of G into $\mathfrak{J}$. The direct computation which follows shows that

$$\pi(f(x)) = \int_G f(x)(\alpha_\pi(x))^*\, dx.$$

If we let g be in $\pi^{-1}(\mathfrak{J})$, then

$$\int_G f(x)(\alpha_\pi(x))^* dx = \left(\int_G f(x)\pi(g_{x^{-1}})dx\right)(\pi(g))^{-1} = \pi(f*g)(\pi(g))^{-1} = \pi(f).$$

Hence there is a 1–1 *correspondence between* $\dagger_*A_2$*-epimorphisms* π *of* A_3 *onto* A_2 *and unitary representatives* α_π *of* G *into* $\mathfrak{J}$.

Let $\bar{G}$ denote the group of all such unitary representations $\alpha_\pi(x)$. The compact-open and weak* topologies (for mappings of A_3 into A_2) are identical for $\bar{G}$. In general, $\bar{G}$ is not locally compact.

The proofs of the last two statements are straightforward and are therefore omitted.

If M_2 in $\mathfrak{M}_2$ is fixed, then $(\alpha_\pi(x))^+(M_2)$, as a function on G is a member of G^+. Hence, for each M_2 in $\mathfrak{M}_2$, there is an epimorphism

$$E_{M_2}: \bar{G} \to G^+,$$

given by:

$$E_{M_2}(\alpha_\pi(x)) = (\alpha_\pi(x))^+(M_2).$$

If π is fixed, then $(\alpha_\pi(x))^+(M_2)$ is a G^+-valued function on $\mathfrak{M}_2$, and actually $(\alpha_\pi(x))^+(M_2)$ is in $C(\mathfrak{M}_2, G^+)$.

If A_2 and A_2^+ are isomorphic, if $A_2^+ = C(\mathfrak{M}_2)$, that is, if A_2 and $C(\mathfrak{M}_2)$ are equivalent, and if β is in $C(\mathfrak{M}_2, G^+)$, define π_β in A_3 by:

$$\pi_\beta(f(x)) = \left(\int_G f(x)\overline{\beta(x; M_2)}\, dx\right)^+ (M_2).$$

Then

$$(\alpha_{\pi_\beta}(x))^+(M_2) = \beta(x; M_2).$$

We have thus far shown that there is a natural mapping γ of $\bar{G}$ into $C(\mathfrak{M}_2, G^+)$ given by

$$\gamma(\alpha_\pi(x)) = \alpha_\pi(x)^+(M_2)$$

and that if A_2 and $C(\mathfrak{M}_2)$ are equivalent, then the natural mapping γ carries $\bar{G}$ *onto* $C(\mathfrak{M}_2, \bar{G}^+)$.

Before stating the next theorem we shall require the following discussion.

If $\alpha(x)$ is in G^+, then for each x there is a unique real number $\beta(x)$, $0 \leqslant \beta(x) < 2\pi$ such that $\alpha(x) = \exp(i\beta(x))$. For example, if G is the circle group (the reals reduced modulo 2π), and if $\alpha(x)$ is in G^+, then there is an integer n such that $\alpha(x) = \exp(i\{nx\})$ where $\{nx\}$ is the residue of nx modulo 2π. Although $\exp(i\{x\})$ is a character in this case, $\exp(i\{\frac{1}{2}\{x\}\})$ is not (since, for example, $\exp(i\{\frac{1}{2}\{2\pi - y\}\}) = \exp(i(\frac{1}{2})(2\pi - y)) \to \exp(i\pi) = -1$ as $y \downarrow 0$, whereas $\exp(i\{\frac{1}{2}\{2\pi - y\}\})$ should approach 1 as $y \downarrow 0$ if $\exp(i\{\frac{1}{2}\{2\pi - y\}\})$ is a character). Hence, in general, even if $\exp(i\beta(x))$ is a character, $\exp(i\{s\beta(x)\})$ is not necessarily a character for all real s.

On the other hand, if G is the additive group of real numbers, and $\alpha(x)$ is in G^+, then there is a real number t such that $\alpha(x) = \exp(i\{tx\})$, $\{tx\}$ the residue of tx modulo 2π. In this case, for any real s, $\exp(i\{s\{tx\}\})$ is again a character.

If a group G has the property that $\exp(i\beta(x))$ is a character implies $\exp(i\{s\beta(x)\})$ is a character, for all real s, we shall call G *real-closed.*

THEOREM 1. *If A_2 is semisimple γ is 1–1; the converse is false. If A_2 and $C(\mathfrak{M}_2)$ are equivalent and G is real-closed, then $\gamma: \bar{G} \to C(\mathfrak{M}_2, G^+)$ is an isomorphism and conversely, if G is real-closed and γ is an isomorphism, then $A_2 = C(\mathfrak{M}_2)$.*

Proof. Assume A_2 is a semisimple and assume

$$\gamma(\alpha_{\pi_1}) = \gamma(\alpha_{\pi_2}).$$

If

$$\alpha_{\pi_1} \neq \alpha_{\pi_2}, \quad \text{then} \quad \pi_1 \neq \pi_2$$

and there is an f in A such that $\pi_1(f) = a_1 \neq a_2 = \pi_2(f)$. But

$$(a_1 - a_2)^+(M_2) = \left(\int_G f(x)(\alpha_{\pi_1}(x) - \alpha_{\pi_2}(x))^* dx\right)^+ (M_2) = 0$$

for all M_2, a contradiction of the semisimplicity of A_2.

Assume A_2 has a radical, R_2. Then R_2 is a non-trivial group R_2° relative to the multiplication: $r_1 \, o \, r_2 = r_1 + r_2 - r_1r_2$.

Now

$$\gamma(\alpha_{\pi_1}) = \gamma(\alpha_{\pi_2}) \text{ if, and only if, } \alpha_{\pi_1}(x) - \alpha_{\pi_2}(x)$$

is in R_2 for all x. But

$$\alpha_{\pi_1}(x) - \alpha_{\pi_2}(x) \in R_2 \text{ for all } x \text{ if, and only if, } 1 - \alpha_{\pi_1}(x)^*\alpha_{\pi_2}(x) = r(x) \in R_2$$

for all x. Clearly $r(x)$ is a representation of G into R_2° (R_2 as a group *re o*). Hence γ is not 1–1 if, and only if, there is a non-trivial representation $r(x)$ of G into R_2°.

However, R_2° contains no elements (different from 0) of finite order. For

$$\overbrace{(r)o(r)o\ldots o(r)}^{n} \equiv r^{n(o)} = 1 - (1-r)^n.$$

If $r^{n(o)} = 0$, r in R_2°, $r \neq 0$, then $(-1)^{N+1}r^N = P_N$ where P_N is a polynomial of degree $N-n$ in n, with coefficients which are polynomials of degree not more than $n-1$ in r. Thus $||r^N|| > n^{N-n}||Q_o||$, where

$$Q_o = \sum_{k=1}^{n-1} (-1)^{k+1}{}_nC_k r^k.$$

Hence $||r^N||^{(1/N)} > n^{(1-n/N)}||Q_o||^{(1/N)} \to n$, as $N \to \infty$, a contradiction of the fact that r is in R_2.

Thus if G has only elements of finite order, $r(x)$ cannot be non-trivial. On the other hand, if $G = R_2^\circ$ (with discrete topology), then $r(x) \equiv x$ serves. Hence the monomorphy of γ depends both on the presence or absence of the radical in A_2 and on the nature of G.

If A_2 and $C(\mathfrak{M}_2)$ are equivalent, then, of course, A_2 is semisimple and hence γ is 1–1, and, as we have shown, an epimorphism. Thus γ is an isomorphism.

On the other hand, if γ is an isomorphism and if A_2^+ is a proper subset of $C(\mathfrak{M}_2)$ let $z(M_2)$ be in the complement of A_2^+. Let $z = u + iv$. Then one of u, v is not in A_2^+, whence we may assume z is real-valued. Since 1 is in A_2^+, A_2^+ contains all constants and hence for some constant c, $z(M_2) + c > 0$,

all M_2. Hence we assume for any $w, 0 < w < 1$, there is a $z(M_2)$ in $C(\mathfrak{M}_2) - A_2^+$ such that

$$w = \inf \{ z(M_2) | M_2 \text{ in } \mathfrak{M}_2 \} < \sup \{ z(M_2) | M_2 \text{ in } \mathfrak{M}_2 \} = 1.$$

Choose $g(x)$ in $L^1(G)$ so that: $g = 0$ outside some compact neighbourhood N of the identity e in G; $g(x) \geq 0$; $g^+(\alpha) \geq 0$; $||g||_1 = 1$; $g^+(\alpha)$ takes on at least three values. Clearly $1 = g^+(e^+) \leq ||g^+||_\infty \leq ||g||_1 = 1$. Let $g^+(\alpha_0) = w \neq 0, 1$. Then $0 < w < 1$ and we now assume $z(M_2)$ and w are related as indicated earlier. If $\alpha_0(x) = \exp(i\beta_0(x))$, let

$$h(s) = \int_G g(x) \exp(i \{s\beta_0(x)\}) \, dx = \int_G g(x) \exp(is\beta_0(x)) \, dx.$$

Then we see that for real s:

(a) $h(s)$ is in $C^\infty(-\infty, \infty)$;

(b) $h(s)$ is real;

(c) $|h^{(n)}(s)| \leq K^n$ *where* $K = \sup \{ |\beta_0(x)| | x \text{ in } N \}$.

Hence $h(s)$ is entire. Since $h(0) = 1$, $h(1) = w < 1$, we see $h(s)$ is not constant. Hence there is an interval (s', s''), $0 < s' < s'' < 1$ where $h'(s) < 0$, and on (s', s'') $h(s)$ has a continuous real-valued inverse: $s = h^{-1}(t)$, where $h(s'') = t'' < t < t' = h(s')$. Let $y(M_2) = az(M_2) + b$ be such that $t'' < y(M_2) < t'$. Then $y(M_2)$ is not in A_2^+ and if $k(M_2) = h^{-1}(y(M_2))$, then $k(M_2)$ is in $C(\mathfrak{M}_2)$, and $k(M_2)$ is real-valued. For $f(x) = g(x)e_2$ (e_2 the identity of A_2) consider

$$\int_G f(x) \exp(ik(M_2)\beta_0(x)) \, dx = \int_G f(x) \exp(i \{k(M_2)\beta_0(x)\}) \, dx$$
$$= h(k(M_2)) = y(M_2).$$

Clearly $\exp(i\{k(M_2)\beta_0(x)\})$ is in $C(\mathfrak{M}_2, G^+)$ and by hypothesis (G is real-closed) there is an α in G such that $\gamma(\alpha_\pi) = \exp(i\{k(M_2)\beta_0(x)\})$. But then $(\pi(f))^+(M_2) = y(M_2)$, contradicting the fact that $y(M_2)$ is not in A_2^+. The proof of the theorem is complete.

4. Miscellany. If G is locally compact abelian, $A_1 = L^1(G)$, and if A_2 has an involution (A_2 is assumed to have no identity and is not assumed to be commutative) easily verified extensions of the above read as follows:

1. *Suppose A_2 is extended to A_{2e} by the adjunction of an identity. Let A_{3e} be the completion of the tensor product of A_1 and A_{2e}. Call an epimorphism $\pi: A_3 \to A_2$ extendable if there is an epimorphism $\pi_e: A_{3e} \to A_{2e}$ which coincides with π on A_3 (naturally embedded in A_{3e}). Then the extendable epimorphisms $\pi: A_3 \to A_2$ are in $1-1$ correspondence with the unitary representations of G into the multiplicative group of A_{2e}.*

2. *If $v(x)$ is a continuous homomorphism*

$$v: G \to G^*_{A_2}$$

(the multiplicative group of A_2 relative to o), then the formula:

$$\pi(f) = \int_G f(x)dx - \int_G f(x)(v(x))^*dx$$

defines an extendable epimorphism π: $A_3 \to A_2$. The extension of π is given by the formula:

$$\pi_e(\lambda(x)e + f(x)) = \int_G (\lambda(x)e + f(x))(e - (v(x))^*)dx.$$

More generally, if $a(x)$ is a numerical function, $a_2(x)$ an A_2-valued function such that $(1 - a(x))(e - a_2(x)) = u(x)$ is a unitary representation of G into the multiplicative group of A_{2e}, then the formula

$$\pi(f) = \int_G f(x)(1 - a(x))dx - \int_G f(x)(a_2(x))^*dx$$

defines an extendable epimorphism $\pi: A_3 \to A_2$. The extension π_e is given by the formula

$$\pi_e(\lambda(x)e + f(x)) = \int_G (\lambda(x)e + f(x))(u(x))^*dx.$$

Conversely, an extendable epimorphism $\pi: A_3 \to A_2$ serves to define two functions $a(x)$, $a_2(x)$ such that $(1 - a(x))(e - a_2(x)) = u(x)$ is a unitary representation of G into the multiplicative group of A_{2e}.

The last result stems from defining T_x on A_{2e} as follows: If a_{2e} is in A_{2e}, and $a_{2e} = \pi_e(\lambda(x)e + f(x))$ then let $T_x(a_{2e}) = \pi_e(\lambda e + f) - \pi_e(\lambda_x e + f_x)$. This definition of T_x is $(\lambda e + f)$-free. T_x satisfies the classic *criterion*:

$$T_x(ab) = (T_x a)b$$

for membership in A_{2e} considered as a subalgebra of the ring $\mathbf{E}(A_e)$ of endomorphisms of A_e. Hence $T_x = a(x)e + a_2(x)$ and the verification of the result follows immediately.

Remark. The *criterion* mentioned above is not valid for algebras having no identity. For example, if $A_2 = L^1(-\infty, \infty)$, $Tf = f_x$, then $T(f^*g) = (Tf)^*g$. But there is no h in A_2 such that $Tf = h^*f$, as is well known.

The standard techniques also show that $a(x)e + a_2(x) = \lim_u \pi_e(u_x e)$ for any approximate identity $\{u\}$ in A.

References

0. N. Bourbaki, *Éléments de mathématique, VII, Première partie, Les Structures fondementales de l'analyse, livre II, Algèbre, chapitre III, Algèbre multilineaire*, Act. Sci. et Indust., *1044* (Paris, 1948), 30–8.

1. A. Hausner, *Abstract 493*, Bull. Amer. Math. Soc. (July, 1956), 383.

2. ——— Proc. Amer. Math. Soc., *8* (1957), 246–9.

3. ——— *The Tauberian heorem for group algebras of vector-valued functions*, Pacific J. Math., *7* (1957), 1603–10.

4. G. P. Johnson, *Abstract 458*, Bull. Amer. Math. Soc. (July, 1956), 366.

5. ——— To appear in Trans. Amer. Math. Soc.

6. R. Schatten, *A theory of cross-spaces* (Princeton, 1950).

7. I. Segal, *The group algebra of a locally compact group*, Trans. Amer. Math. Soc., *61* (1947), 69–105.

8. A. B. Willcox, *Note on certain group algebras*, Proc. Amer. Math. Soc., *7* (1956), 874–9.

University of Minnesota

A CLASS OF SOLVABLE GROUPS

DANIEL GORENSTEIN AND I. N. HERSTEIN

1. Introduction. Numerous studies have been made of groups, especially of finite groups, G which have a representation in the form AB, where A and B are subgroups of G. The form of these results is to determine various group-theoretic properties of G, for example, solvability, from other group-theoretic properties of the subgroups A and B.

More recently the structure of finite groups G which have a representation in the form ABA, where A and B are subgroups of G, has been investigated. In an unpublished paper, Herstein and Kaplansky **(2)** have shown that if A and B are both cyclic, and at least one of them is of prime order, then G is solvable. Also Gorenstein **(1)** has completely characterized ABA groups in which every element is either in A or has a *unique* representation in the form aba', where a, a' are in A, and $b \neq 1$ is in B.

In this paper we shall analyse groups of the form ABA in which A and B are cyclic of relatively prime order. The techniques and methods used borrow heavily from those used in the aforementioned paper of Herstein and Kaplansky. The authors became interested in the structure of ABA groups as an outgrowth of problems they considered while at a conference held at Bowdoin College in the summer of 1957 under the auspices of the Cambridge Research Center of the United States Air Force.

In the body of the paper we shall use the following notation: If H is a subgroup of G, $o(H)$, $i(H)$, and $N(H)$ will denote respectively the order of H, the index of H in G, and the normalizer of H in G.

2. Two preliminary lemmas. We shall need a result on the transfer homomorphism which is a slight extension of a result of Grün **(3**, p. 143**)**; in fact, the result is essentially contained in Grün's, but for the sake of completeness we present it here.

LEMMA 1. *Let G be a finite group, and A an Abelian subgroup of G for which $(o(A), i(A)) = 1$. Then the transfer of G into A maps the intersection of A with the centre of its normalizer onto itself.*

Proof. Since $(o(A), i(A)) = 1$ it is clear that for any $p|o(A)$ the p-Sylow subgroup of A is a p-Sylow subgroup of G.

We first contend that if A_1 is an Abelian subgroup of G and $o(A_1) = o(A)$, then A_1 is a conjugate of A. Let $S_p \neq (1)$ be the p-Sylow subgroup of A_1.

Received June 16, 1958. The work of the second author was supported in part by OOR, Contract No. ORDOR-LO-P-2042/A11472.

Since these are p-Sylow subgroups of G, there is a $y \in G$ such that $S_p = yS_p'y^{-1}$. If we replace A_1 by yA_1y^{-1} we may, without loss of generality, assume that $S_p \subset A$ and $S_p \subset A_1$.

If $N(S_p) \neq G$, our contention follows by induction and from the fact that both A and A_1, being Abelian, are contained in $N(S_p)$. If, on the other hand, $N(S_p) = G$, we use induction on $\bar{G} = G/S_p$ to conclude that $\bar{A}$, $\bar{A}_1$, the images of A and A_1 in $\bar{G}$ are conjugate in $\bar{G}$. Since both A and A_1 contain S_p, their conjugacy in G follows at once.

From this, and the usual argument made on the centres of Sylow subgroups, we can say that if two elements of A are conjugate in G then they are already conjugate in $N(A)$.

We are now able to prove the lemma. For let a_1 be an element in the intersection of A with the centre of $N(A)$, we compute the transfer, τ, on a_1. Since A is Abelian,

$$\tau(a_1) = \prod_{i=1}^{r} x_i a_1^{f_i} x_i^{-1} \quad \text{where} \quad \sum_{i=1}^{r} f_i = i(A) \quad \text{and} \quad x_i a_1^{f_i} x^{-1} \in A.$$

However, since

$$a_1^{f_i} \quad \text{and} \quad x_i a_1^{f_i} x_i^{-1}$$

are conjugate in G and are in A they are conjugate in $N(A)$; since a_1 is in the centre of $N(A)$ they must be equal. Thus

$$\tau(a_1) = \prod_{i=1}^{r} a_1^{f_i} = a_1^{i(A)};$$

and since $(o(A), i(A)) = 1$, the lemma follows.

The second result we shall need is contained in the following lemma.

LEMMA 2. *Suppose a finite group G admits an automorphism α of order h such that every element of G can be expressed in the form $\alpha^i(b^j)$ for some fixed element b of G of order k. If $(h, k) = 1$, then G is either Abelian or is the direct product of an Abelian group of odd order with the quaternion group of order* 8. *If α leaves only the identity element of G fixed, then G is Abelian.*

Proof. We proceed by induction on the order of G.

Suppose, first, that α leaves some element, $\neq 1$, of G fixed. Then for some $e, t, \alpha(\alpha^e(b^t)) = \alpha^e(b^t), b^t \neq 1$. Thus $\alpha(b^t) = b^t$. But then for all i, j

$$\alpha^i(b^j) \,.\, b^t = \alpha^i(b^j)\alpha^i(b^t) = \alpha^i(b^{j+t}) = b^t\alpha^i(b^j),$$

and so b^t is in Z, the centre of G.

Since the order of every element of G is a divisor of the order of b, then for any prime p, $p \mid o(G)$ implies $p \mid k$. We consider the cases when $p \nmid k/t$ and $p \mid k/t$ separately.

Suppose first that $p \nmid k/t$. Let $\bar{G} = G/(b^t)$ and let $\bar{\alpha}$ be the automorphism induced by α on $\bar{G}$. If $\bar{b}$ is the image of b in $\bar{G}$, then every element of $\bar{G}$ is clearly of the form $\bar{\alpha}^i(\bar{b}^j)$. By our induction hypothesis the p-Sylow subgroup

$\bar{S}_p$ of $\bar{G}$ is normal in $\bar{G}$. Thus the inverse image of $\bar{S}_p$ in $\bar{G}$, is normal in G and is of the form $S_p\cdot(b^t)$ for some p-Sylow subgroup of G. If $s \in S_p$, $x \in G$, then $xsx^{-1} = b^{tj}s_1$, $s_1 \in S_p$; since the order k/t of b^t, is relatively prime to p, this implies $b^{tj} = 1$, and so $xS_px^{-1} = S_p$, so S_p is normal in G.

Furthermore

$$\bar{S}_p = \frac{S_p\cdot(b^t)}{(b^t)} \cong \frac{S_p}{S_p \cap (b^t)} = S_p;$$

by induction $\bar{S}_p$, and hence S_p, is either Abelian or isomorphic to the quaternion group of order 8.

If, on the other hand, $p|k/t$ then $b^{k/p}$ is in Z, and being of prime-power order, must be in all p-Sylow subgroups of G. By induction, $\bar{S}_p$, the p-Sylow subgroup of $\bar{G} = G/(b^{k/p})$ is normal in $\bar{G}$, so its inverse image, S_p, must be normal in G. Thus S_p contains all elements of G whose order is a power of p. We claim S_p contains a unique subgroup of order p. For if $\alpha^i(b^j)$ is of order p, then j is a multiple of k/p, so $\alpha^i(b^j) = b^j$ since $\alpha(b^{k/p}) = b^{k/p}$; thus the only subgroup of order p is $(b^{k/p})$. It is well known that a group of prime-power order having only one subgroup of order p is cyclic if p is odd and is either cyclic or a generalized quaternion group of order 2^n if $p = 2$. Now S_2 is a normal subgroup of G invariant under α. Since by assumption $(h, k) = 1$ and since $2|k$ if $S_2 \neq 1$, we conclude that α is of odd order. If α reduces to the identity on S_2, S_2 is cyclic. Hence if S_2 is isomorphic to the generalized quaternion group, α has odd order on S_2. But for $n > 3$ the automorphism group of a generalized quaternion group of order 2^n is of order 2^{n-1}. Thus the only possibility in our case is $n = 3$, and S_2 is isomorphic to the quaternion group of order 8.

There remains the case when α leaves no element of G, other than 1, fixed. In this situation it is known that for each prime p, α must leave some p-Sylow subgroup, say, S_p, fixed. Since an element of S_p is of the form $\alpha^i(b^j)$, it follows readily that S_p consists of all the elements of G whose order is a power of p. S_p is then the unique p-Sylow subgroup of G and so is normal in G, and G is the direct product of its Sylow subgroups. We still must show that S_p is either Abelian or the quaternion group of order 8. Thus we may, without loss of generality, assume that $S_p = G$.

Suppose then that $k = p^s$. We compute the number of elements in G. Let r_i be the least positive integer such that

$$\alpha^{r_i}(b^{p^i}) \in (b^{p^i}).$$

It is clear that the number of elements in G, of order exactly p^{s-i}, is $r_i(p^{s-i} - p^{s-i-1})$, $i < s$, and so

$$o(G) = p^n = r_0(p^s - p^{s-1}) + r_1(p^{s-1} - p^{s-2}) + \dots + r_{s-1}(p-1) + 1. \tag{2.1}$$

However, the elements of order p in the centre of G form a characteristic subgroup of G, and so the elements of the form

$$\alpha^i(b^{p^{s-1}j})$$

(that is, all the elements of order p), form a subgroup of G (in Z) of order p^m, containing $r_{s-1}(p-1)+1$ elements. So

$$p^m = r_{s-1}(p-1)+1. \tag{2.2}$$

Combining (2.1) and (2.2) we have

$$p^n - p^m = r_0(p^s - p^{s-1}) + \ldots + r_{s-2}(p^2 - p).$$

If $m > 1$, then p^2 divides the left-hand side, and so must divide the right-hand side; but then $p|r_{s-2}$. Since $r_{s-2}|h$, and $1 = (h, k) = (h, p^s)$, this is impossible. So $m = 1$, and we can conclude that G has exactly one subgroup of order p. We conclude, as above, that G is either cyclic or isomorphic to the quaternion group of order 8.

The final statement of the lemma follows at once from the fact that the quaternion group has a unique element of order 2 and hence each of its automorphisms leaves this element fixed.

3. The case $N(A) = A$. In this section we shall prove the following result concerning the structure of ABA groups:

THEOREM 1. *Let G be an ABA group, in which A and B are cyclic subgroups of relatively prime orders h and k respectively. Then if A is its own normalizer in G, G contains a normal subgroup T with $A \cap T = 1$. Furthermore T is either Abelian or the direct product of an Abelian group of odd order with the quaternion group of order 8. In particular, G is solvable, and of order hkw, where $w|k^\nu$ for some integer ν.*

Proof. We shall prove first that the Sylow subgroups of A are, in fact, Sylow subgroups of G. The proof is by induction on the order of G.

Let S_p be a p-Sylow subgroup of A. Since A is Abelian, $N(S_p) \supset A$. If $x = a_1b_1a_2 \in N(S_p)$ with $b_1 \in B$, $a_1, a_2 \in A$, then clearly $b_1 \in N(S_p) \cap B$. If $B_1 = B \cap N(S_p)$, then obviously $N(S_p)$ is of the form AB_1A; thus if $N(S_p)$ is a proper subgroup of G, it follows by induction that the order of $N(S_p)$ is hk_1w_1, where $k_1 = o(B_1)$ and $w_1|k_1^\nu$ for some integer ν. Since $(h, k_1) = 1$, we see then that S_p is a p-Sylow subgroup of $N(S_p)$. But S_p must then be a p-Sylow subgroup of G, since the normalizer of a proper subgroup of a p-group is always a strictly larger subgroup.

On the other hand, if $N(S_p) = G$, then S_p is normal in G, and we consider $\bar{G} = G/S_p = \bar{A}\bar{B}\bar{A}$, where $\bar{A}$, $\bar{B}$ are the images of A and B. Furthermore $N(\bar{A}) = \bar{A}$; for $N(\bar{A}) > \bar{A}$ would clearly imply $N(A) > A$ since S_p is contained in A. Hence we can apply our induction hypothesis to $\bar{G}$, and we obtain $o(G) = \bar{h}k\bar{w}$, where $\bar{w}|k^\nu$ and $\bar{h} = o(\bar{A})$. Thus $p \nmid o(\bar{G})$, and so S_p is a Sylow subgroup of G.

Since this holds for every $p|h$, it follows that the order and index of A are relatively prime. Since A is Abelian, we may apply Lemma 1 to conclude that the transfer τ of G into A maps the intersection of A with the centre

of its normalizer onto itself. But by assumption, A is the centre of its normalizer, and so τ maps G homomorphically *onto* A.

Let T be the kernel of τ; since τ maps A onto itself, $A \cap T = 1$; since the order of B is relatively prime to the order of A, $B \subset T$. If a, b are generators of A, B respectively, it is clear that T consists precisely of the elements of G of the form $a^i b^j a^{-i}$, where i, j are arbitrary. Now the mapping α defined on T by $\alpha(x) = axa^{-1}$ is an automorphism of T, and every element of T is of the form $\alpha^i(b^j)$. Therefore by Lemma 2, T is of the form stated in the theorem.

Since every element of T, being of the form $\alpha^i(b^j)$, has order a divisor of k, $o(T) = kw$ where $w | k^\nu$ for some integer ν. Since $G/T = A$, and A is cyclic of order h, G is solvable of order hkw. This completes the induction and the proof of the theorem.

COROLLARY. *A Sylow subgroup of G is either Abelian or isomorphic to the quaternion group of order* 8.

4. The main theorem.

THEOREM 2. *Let G be an ABA group in which A and B are cyclic subgroups of relatively prime orders h and k respectively. Then*

1. *G is solvable.*
2. *The p-Sylow subgroups of G, for odd p, are Abelian;*
3. *The 2-Sylow subgroup of G is either Abelian or isomorphic to the quaternion group of order* 8.
4. *The order of G is hkw, where $w | k^\nu$ for some integer ν.*

Proof. If $N(A) = A$, the theorem follows immediately from Theorem 1 and its corollary. We may therefore assume that $N(A) > A$. If a, b denote, as above, generators of A and B, there is an element of the form $a^i b^j a^e$ in $N(A)$ with $b^j \neq 1$, whence b^j itself is in $N(A)$. Let r be the least positive integer such that $b^r \in N(A)$. Then $r | k$, and we have for some integer λ

$$b^r a b^{-r} = a^\lambda, \qquad \text{where } \lambda^{k/r} \equiv 1 \pmod{h}. \tag{4.1}$$

Let p be a prime dividing k/r and define γ_p as the least multiple of r such that k/r is a power of p. Set

$$B_p = (b^{\gamma_p}).$$

In the first part of the proof we shall establish the following statement:

The normalizer $N(B_p)$ of B_p is of the form $A_p B A_p$ for a suitable subgroup A_p of A, and furthermore B_p is in the centre of $N(B_p)$.

We shall need one preliminary result. Since $r | \gamma_p$, we have

$$b^{\gamma_p} a b^{-\gamma_p} = a^{\lambda_p}, \quad \lambda_p^{k/\gamma_p} \equiv 1 \pmod{h} \tag{4.2}$$

for some integer λ_p. Let $u_p = (\lambda_p - 1, h)$. We assert that

$$\left(u_p, \frac{h}{u_p}\right) = 1.$$

For suppose a prime $q \mid (\lambda_p - 1)$. It is sufficient to show that if $q^\epsilon \mid h$, then $q^\epsilon \mid (\lambda_p - 1)$. Since $k/\gamma_p = p^s$ for some integer s, (4.2) implies that

$$\lambda_p^{p^s} \equiv 1 \qquad (\mathrm{mod}\ h)$$

and hence

$$\lambda_p^{p^s} \equiv 1 \qquad (\mathrm{mod}\ q^\epsilon).$$

Write $\lambda_p = 1 + xq^\delta$, where $(x, q) = 1$. Then $(1 + xq^\delta)^{p^s} \equiv 1 \pmod{q^\epsilon}$ and so $p^s xq^\delta + yq^{2\delta} \equiv 0 \pmod{q^\epsilon}$ for some integer y. Since $p \mid k$ and $q \mid h$, and $(h, k) = 1$, $p \neq q$ and hence $\delta \geqq \epsilon$.

We now return to $N(B_p)$. Suppose that $a^i b^j a^e \in N(B_p)$. Then

$$a^i b^j a^e b^{\gamma_p} a^{-e} b^{-j} a^{-i} = b^{\gamma_p m}$$

for some integer m.

Applying (4.2) to this relation, we obtain

$$a^i b^j a^{e(1-\lambda_p)} b^{-j} a^{-i\lambda_p} = b^{\gamma_p(m-1)}. \tag{4.3}$$

Suppose first that

$$\frac{h}{u_p} \,\Big|\, e.$$

Since $u_p \mid (\lambda_p - 1)$, (4.3) reduces to

$$a^{i(1-\lambda_p)} = b^{\gamma_p(m-1)}$$

and their common value is 1, since a and b have relatively prime order. But $u_p = (1 - \lambda_p, h)$, and so

$$\frac{h}{u_p} \,\Big|\, i;$$

moreover,

$$m \equiv 1 \left(\mathrm{mod}\ \frac{k}{\gamma_p}\right),$$

and hence $a^i b^j a^e$ commutes with b^{γ_p}.

Conversely, every element of the form

$$a^{\frac{h}{u_p} i} b^j a^{\frac{h}{u_p} e}$$

is in $N(B_p)$ and commutes with b^{γ_p}. To complete the proof of our assertion, we shall show that every element of $N(B_p)$ is of this form. We have just shown this to be the case if

$$\frac{h}{u_p} \,\Big|\, e.$$

Suppose, on the other hand, that $e(1-\lambda_p) \not\equiv 0 \pmod{h}$. Then (4.3) yields the relation

$$b^j a^{e(1-\lambda_p)} b^{-j} = a^{-i} b^{\gamma_p(m-1)} a^{i\gamma_p},$$

and hence $b^j a^{e(1-\lambda_p)} b^{-j}$ is in $N(A)$. But this element has order dividing h; since $(h, k) = 1$, all the elements of $N(A)$ of order dividing h are already in A. Thus

$$b^j a^{e(1-\lambda_p)} b^{-j} = a^{e(1-\lambda_p)\rho} \tag{4.4}$$

for some integer ρ.

Using (4.4), we can rewrite the element

$$a^i b^j a^e = a^i b^j a^{-xe(1-\lambda_p)} a^{e+xe(1-\lambda_p)}$$

as

$$a^{i-xe(1-\lambda_p)\rho} b^j a^{e+xe(1-\lambda_p)}.$$

Since

$$\left(1-\lambda_p, \frac{h}{u_p}\right) = 1,$$

we can find an integer x such that

$$e + xe(1-\lambda_p) \equiv 0 \left(\operatorname{mod} \frac{h}{u_p}\right).$$

If, for this x, we set $i' = i - xe(1-\lambda_p)\rho$ and $e' = e + xe(1-\lambda_p)$, we have $a^i b^j a^e = a^{i'} b^{j'} a^{e'}$, where

$$\left.\frac{h}{u_p}\right| e';$$

and hence

$$a^i b^j a^e{}_h = a^{m\frac{h}{u_p}}\, b^j\, a^{n\frac{h}{u_p}}.$$

If

$$A_p = (a^{u_p}),$$

we have thus proved that $N(B_p) = A_p B A_p$, and that B_p is in the centre of $N(B_p)$.

5. Continuation of the proof. The proof now proceeds by induction on the order of G, but we add the following statement to our induction hypothesis: if $p|k$ then some p-Sylow subgroup of G consists of all the elements of the form $a^{si}b^{tj}a^{-si}$ for suitable integers s, t where i, j are arbitrary.

There are three cases to consider, which we take up in succession.

Case 1. $N(B_p) = G$. In this case B_p is in the centre of G, and we define $\bar{G} = G/B_p = \bar{A}\bar{B}\bar{A}$, where $\bar{A}$ has order h and $\bar{B}$ has order γ_p. By induction, $\bar{G}$ is solvable of order $h\gamma_p w$, where $w|\gamma_p'$; so G is solvable and its order is

$$(h\gamma_p w)\frac{k}{\gamma_p} = hkw$$

where $w|\gamma_p{}^r|k^r$.

Hence the order of A is relatively prime to its index in G. Thus the Sylow subgroups of G, for any prime dividing h, are cyclic.

Furthermore, the Sylow subgroups $\bar{S}_q$ of $\bar{G}$ for primes q which divide γ_p are of the form $\{\bar{a}^{st}\bar{b}^{tj}\bar{a}^{-st}\}$ for suitable integers s and t. If $q \neq p$, then it follows as in the proof of Lemma 2 that the elements

$$\{a^{st}b^{t\frac{k}{\gamma_p}j}a^{-st}\}$$

form a q-Sylow subgroup of G which maps isomorphically on $\bar{S}_q$. If $q = p$, it follows again as in the proof of Lemma 2 that the complete inverse image of a suitable p-Sylow subgroup $\bar{S}_p$ of $\bar{G}$ is a p-Sylow subgroup of G and is of the form $\{a^{st}b^{tj}a^{-st}\}$ for suitable s, t. Thus for each prime p dividing k a p-Sylow subgroup S_p of G is of the required form. If α denotes the automorphism of S_p defined by $\alpha(x) = a^s x a^{-s}$ for x in S_p, then every element of S_p is of the form $\alpha^t(b^{tj})$. Hence by Lemma 2, S_p is either Abelian or isomorphic to the quaternion group of order 8. Our induction is therefore complete in the case $N(B_p) = G$.

We may therefore assume $N(B_p) < G$.

Case 2. $p \neq 2$. By our induction hypothesis some p-Sylow subgroup S_p of $N(B_p)$ is of the form

$$\{a^{\frac{h}{u_p}\sigma t}b^{\gamma_1 j}a^{-h/u_p\sigma t}\}$$

for some integers σ and γ_1, and is Abelian. Since B_p is in the centre of $N(B_p)$, $B_p \subset S_p$ and hence $\gamma_1|\gamma_p$. We shall prove first that S_p is cyclic. Suppose b^{γ_1} has order p^δ. Then clearly S_p is Abelian of type $(p^\delta, p^\delta, \ldots, p^\delta)$. But since

$$a^{h/u_p}$$

commutes with b^{γ_p}, B_p is the only subgroup of its order in S_p, and hence S_p is cyclic.

Since S_p is cyclic, B_p is a characteristic subgroup of S_p, and so $N(S_p) \subset N(B_p)$. If S_p were not a p-Sylow subgroup of G, its normalizer would contain a strictly larger p-group than S_p. But since S_p is a p-Sylow subgroup of $N(B_p)$ and since $N(S_p) \subset N(B_p)$, it must be that S_p is in fact a p-Sylow subgroup of G.

But now by a theorem of Grün (or by Lemma 1), the transfer τ of G into the Abelian Sylow subgroup S_p maps G onto the intersection of S_p with the centre of its normalizer. But B_p is contained in the centre of its normalizer. Thus τ maps G homomorphically on (b^{γ_2}) where $\gamma_1|\gamma_2|\gamma_p$. Since the order of A is relatively prime to that of B, A is contained in the kernel H of τ. Also H contains some proper subgroup B^* of B, and hence H is of the form AB^*A. Since G/H has order k/γ_2, B^* is of order γ_2. Thus for any prime different from p, the Sylow subgroups of H are Sylow subgroups of G, and hence by induction

are of the required form. As we have already seen S_p itself is cyclic and of the required form. By induction H is solvable and $o(H) = h\gamma_2 w$ where $w|\gamma_2^\nu$ for some integer ν. Since H and G/H are solvable, G is solvable and the order of G is $o(G/H)o(H)$. Thus $o(G) = (h\gamma_2 w)k/\gamma_2 = hkw$, where $w|\gamma_2^\nu|k^\nu$. It follows at once that the Sylow subgroups of G, for any prime dividing h, is cyclic. The proof is complete in this case.

Case 3. $p = 2$. If S_2 is Abelian, or if an odd prime divides k/r, the above proof holds without change. There remains then but one case to consider: namely, when $k/r = 2^s$ and the 2-Sylow subgroup S_2 of $N(B_2)$ is isomorphic to the quaternion group of order 8. In this case $\gamma_p = \gamma_2 = r$ and $B_2 = (b^r)$.

Since the quaternion group has no element of order 8, $8 \nmid k$. On the other hand, suppose $2 \nmid r$. Then $\bar{N} = N(B_2)/B_2$ is of the form $\bar{A}_2\bar{B}\bar{A}_2$, where $\bar{B}$ has order r. Then by our induction hypothesis, $o(\bar{N}) = o(\bar{A}_2)o(\bar{B})\bar{w} = o(\bar{A}_2)r\bar{w}$, where $\bar{w}|r^\nu$, and so $o(\bar{N})$ would be odd. But then $S_2 = B_2$, contrary to our assumption that S_2 is the quaternion group. Hence we must have $2|r$. Since k/r is a power of 2 and $8 \nmid k$, it follows that $r/2$ is odd and k/r is 2.

We are thus reduced to considering the following situation:

$$b^r a b^{-r} = a^\lambda, \quad 2|r, \quad r/2 \text{ is odd}, \quad \text{and } k/r = 2. \tag{5.1}$$

Let $(\lambda - 1, h) = u$. If $u = 1$, then $N(B_2) = B$ (since $N(B_2)$ is of the form $\{a^{h/u\, i}b^j a^{-h/u\, i}\}$). But then again S_2 would be cyclic. So we may assume that $u > 1$.

We may further assume that no subgroup of A is normal in G, for the theorem follows easily by induction in this case. In particular, this implies, as in the proof of Theorem 1, that the order of A is relatively prime to its index.

From (5.1),we have $\lambda^2 - 1 \equiv 0 \pmod h$, and hence $u(\lambda + 1) \equiv 0 \pmod h$. Thus $b^r a^u b^{-r} = a^{u\lambda} = a^{-u}$, and similarly $b^r a^{-u} b^r = a^u$. Now as we have already seen, b^r commutes with $a^{h/u}$. But $N(A)$ is generated by a and b^r, and hence $a^{h/u} \neq 1$ is in the centre of $N(A)$. Since the order of A is relatively prime to its index, it follows from Lemma 1 that the transfer of G into A maps the intersection of A with the center of its normalizer onto itself. Thus $(a^{h/u})$ is mapped onto itself by the transfer map. Hence the kernel H of the transfer of G into A consists of all elements of the form $a^i b^j a^{-i+su}$. Suppose now that $x = a^i b^j a^{-i+su}$ is an element of order 2 in H. Thus $1 = x^2 = a^i b^j a^{-i+su} a^i b^j a^{-i+su}$, whence

$$b^j a^{su} b^j = a^{-su}. \tag{5.2}$$

Conjugating this relation by b^r we obtain $b^j a^{-su} b^j = a^{su}$. Thus $b^j a^{su} b^j b^j a^{-su} b^j = a^{-su} a^{su} = 1$, and so

$$a^{su} b^{2j} a^{-su} = b^{-2j}. \tag{5.3}$$

Thus $a^{2su} b^{2j} a^{-2su} = b^{2j}$. But $(2,h) = 1$, so that we must have $a^{su} b^{2j} a^{-su} = b^{2j}$. Equation (5.3) now yields that $b^{4j} = 1$. Consequently $k|4j$; that is $k/4 = (r/2)|j$ Thus $x^2 = 1$ implies that

$$x = a^i b^{\frac{r}{2}m} a^{-i+su}.$$

Suppose next that $j = r/2$. Then (5.2) becomes $b^{\frac{1}{2}r}a^{eu}b^{\frac{1}{2}r} = a^{-eu}$, so that $b^{\frac{1}{2}r}a^{eu}b^{-\frac{1}{2}r} = a^{-eu}b^{-r} = a^{-eu}b^{r}$ since $k = 2r$; but now the element on the right-hand side of this relation is of order 2, while that on the left-hand side is not, which is a contradiction. Similarly, $j = -\frac{1}{2}r$ is impossible.

We have thus proved that if $x = a^{t}b^{j}a^{-t+eu}$ is an element of order 2 in H, then $j = r$, and so x is of the form $b^{r}a^{e'u}$. Since H is normal in G, and since $b^{r}a^{u}$ is of order 2, $b(b^{r}a^{u})b^{-1}$ is of order 2 and is in H. It follows that $b(b^{r}a^{u})b^{-1} = b^{r}a^{mu}$ for some integer m, and hence $ba^{u}b^{-1} = a^{mu}$. Thus a^{u} generates a normal subgroup of G, in contradiction to our present assumption that no subgroup of A is normal in G. This contradiction completes the proof of the theorem.

References

1. D. Gorenstein, *A Class of Frobenius Groups*, Can. J. Math., *11* (1959), 39–47.
2. I. N. Herstein and I. Kaplansky, *Groups of Cyclic Length Three*, project document No. 13, Summer Mathematical Conference, Bowdoin College, Brunswick, Maine (1957).
3. H. Zassenhaus, *The Theory of Groups*, New York, 1949).

Clark University
and
Cornell University